

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE CIENCIAS BIOLÓGICAS**  
**Departamento de Genética**



**EL GEN “TRANSFORMER-2” DE SCIARA  
(DÍPTERA, NEMATOCERA, SCIARIDAE).**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR**  
**PRESENTADA POR**

**Iker Martín Sánchez**

Bajo la dirección de los doctores

Lucas Sánchez Rodríguez  
María Fernanda Ruiz Lorenzo

**Madrid, 2010**

**ISBN: 978-84-693-9251-5**

**© Iker Martín Sánchez, 2010**

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS

Departamento de Genética



**El gen *transformer-2* de *Sciara*  
(Díptera, Nematocera, Sciaridae)**

Tesis Doctoral presentada por

**IKER MARTÍN SÁNCHEZ**

Madrid, 2010

D. Lucas Sánchez Rodríguez, Profesor de Investigación del Centro de Investigaciones Biológicas (C.S.I.C) y D<sup>a</sup> María Fernanda Ruiz Lorenzo, Investigadora contratada del Centro de Investigaciones Biológicas (C.S.I.C),

CERTIFICAN

Que D. IKER MARTÍN SÁNCHEZ, Licenciado en Ciencias Biológicas por la Universidad Complutense de Madrid, ha realizado bajo nuestra dirección el trabajo correspondiente a su Tesis Doctoral, con el título: El gen *transformer-2* de *Sciara* (Díptera, Nematocera, Sciaridae). Dicho trabajo reúne los requisitos para su presentación.

Madrid, 3 de mayo de 2010

El Director de la Tesis

La Co-directora de la Tesis

Prof. Lucas Sánchez Rodríguez

Dra. María Fernanda Ruiz Lorenzo

El proceso de la determinación sexual se conoce bien en *Drosophila melanogaster*. La caracterización molecular y la organización jerárquica de los genes que participan en este proceso constituyen una base para el aislamiento de genes homólogos en otras especies de insectos y poder estudiar los diferentes tipos de determinación sexual existentes. El gen *transformer-2* (*tra2*) de *Drosophila*, se expresa de manera constitutiva en machos y en hembras. La proteína Tra2 pertenece a la familia de proteínas SR (*Splicing Regulador*) y presenta una dualidad de función actuando como un activador del procesamiento del transcrito primario del gen *dsx* en las hembras y, como un inhibidor del procesamiento de su propio transcrito primario durante la espermatogénesis en el macho adulto.

En esta Tesis doctoral se ha llevado a cabo el aislamiento y la caracterización de un gen en *Sciara ocellaris* y en *Sciara coprophila* homólogo al gen *transformer-2* (*tra2*) encontrado en otros insectos.

El gen *Sotra2* está formado por seis exones y cinco intrones. Se transcribe a lo largo de todo el desarrollo, en machos y hembras, dando lugar a tres productos de distinto tamaño por un procesamiento alternativo no específico de sexo. El transcrito mayoritario *Sotra2-251* codificaría una proteína putativa SoTra2-251 de 251 aminoácidos homóloga a la proteína Tra2 de otros insectos. Los otros dos transcritos, *Sotra2-204* y *Sotra2-130*, codificarían, respectivamente, una proteína Tra2 truncada.

El gen *Sctra2* está formado por seis exones y cinco intrones. Se transcribe a lo largo de todo el desarrollo en ambos sexos dando lugar a cuatro productos de distinto tamaño por procesamiento alternativo no específico de sexo. El transcrito mayoritario *Sctra2-246* codificaría una proteína putativa ScTra2-246 de 246 aminoácidos homóloga a la proteína Tra2 de otros insectos. Los otros tres transcritos, *Sctra2-203*, *Sctra2-103* y *Sctra2-75* codificarían, respectivamente, una proteína putativa Tra2 truncada.

La proteína Tra2 de *Sciara* posee las características que definen a la familia de proteínas SR; esto es, contiene un dominio RRM de unión a ARN, flanqueado por dos dominios RS1 y RS2 ricos en dipéptidos de arginina-serina, y una región "linker" que une el dominio RRM con el RS2.

La comparación de la proteína Tra2 de *Sciara* con la de otros insectos evidencia que el mayor grado de conservación ocurre en el dominio RRM y en la región "linker", mientras que los dominios RS1 y RS2 presentan variabilidad en cuanto al número de aminoácidos que los componen y en cuanto al número de dipéptidos arginina-serina que poseen.

Hemos producido moscas de *Drosophila melanogaster* transgénicas para la proteína SoTra2-251. Hemos determinado que SoTra2-251 es capaz de inducir el procesamiento alternativo específico de hembra del transcrito primario del gen *dsx*, mimetizando la función de la propia proteína Tra2 de *Drosophila*. Sin embargo, este complejo no parece ser tan eficiente como el propio complejo Tra-Tra2 de *D. melanogaster* en inducir ese procesamiento, lo que sugiere la existencia de co-evolución entre las proteínas Tra y Tra2.

<b>1. INTRODUCCIÓN</b>	<b>5</b>
<b>1.1 MECANISMOS DE DETERMINACIÓN SEXUAL EN INSECTOS</b>	<b>6</b>
<b>1.2 DETERMINACIÓN SEXUAL EN <i>Drosophila melanogaster</i></b>	<b>8</b>
<b>1.3 EL GEN tra 2 DE <i>Drosophila melanogaster</i></b>	<b>12</b>
1.3.1 Las proteínas Tra2	13
1.3.2 Función	14
<b>1.4 GENES DE LA DETERMINACIÓN SEXUAL EN OTROS INSECTOS</b>	<b>18</b>
1.4.1 El gen <i>Sex-lethal</i>	18
1.4.2 El gen <i>transformer</i>	19
1.4.3 El gen <i>transformer-2</i>	20
1.4.4 El gen <i>doublesex</i>	22
1.4.5 El gen <i>fruitless</i>	22
1.4.6 El gen <i>intersex</i>	22
<b>1.5 MECANISMOS DE DETERMINACIÓN SEXUAL EN <i>Sciara</i></b>	<b>23</b>
<b>2. OBJETIVOS</b>	<b>27</b>
<b>3. MATERIALES Y MÉTODOS</b>	<b>29</b>
<b>3.1 MATERIALES</b>	<b>30</b>
3.1.1 Productos	30
3.1.2 Medios de cultivo y soluciones	30
3.1.3 Mantenimiento de <i>S. coprophila</i> y <i>S. ocellaris</i>	30
3.1.4 Mantenimiento de <i>Drosophila melanogaster</i>	30
<b>3.2 MÉTODOS</b>	<b>31</b>
<b>3.2.1 TÉCNICAS MOLECULARES DE ADN Y ARN</b>	<b>31</b>
3.2.1.1 Extracción de ADN genómico	31
3.2.1.2 Extracción de ADN plasmídico	31
3.2.1.3 Electroforesis en geles de agarosa	31
3.2.1.4 Amplificación de fragmentos de ADN mediante PCR ( <i>Polymerase Chain Reaction</i> )	32
3.2.1.5 Clonaje de productos de PCR	32
3.2.1.6 Comprobación de colonias positivas por PCR	32
3.2.1.7 Comprobación de colonias positivas mediante "cracking"	33
3.2.1.8 Comprobación de colonias positivas mediante digestión con endonucleasas de restricción	33
3.2.1.9 Extracción de ARN total	34
3.2.1.10 Extracción de ARN poliadenilado	34
3.2.1.11 Retrotranscripción (RT)	34
3.2.1.12 Secuenciación de ADN	34
3.2.1.13 Análisis de las secuencias	35
3.2.1.14 Análisis del posible sitio de inicio de la transcripción	35
3.2.1.15 Diseño de los oligonucleótidos.	35
<b>3.2.2 AISLAMIENTO Y CARACTERIZACIÓN DEL GEN tra2 DE <i>S. ocellaris</i></b>	<b>35</b>
3.2.2.1 RT-PCR con oligonucleótidos degenerados	35
3.2.2.2 RACE (" <i>Rapid Amplification of cDNA Ends</i> ")	36
3.2.2.3 <i>Genome Walker</i> (Paseo Cromosómico)	37
3.2.2.4 PCR de largo alcance (" <i>long PCR</i> ")	37

3.2.2.5 Expresión del gen <i>tra2</i> de <i>S. ocellaris</i> durante el desarrollo mediante RT-PCR	38
3.2.2.6 Expresión del transcrito <i>SoTra2-204</i>	39
3.2.2.7 Expresión del transcrito <i>SoTra2-130</i>	39
3.2.2.8 Cuantificación de la expresión de los transcritos <i>SoTra2-251</i> y <i>SoTra2-130</i> en <i>S. ocellaris</i>	40
<b>3.2.3 AISLAMIENTO Y CARACTERIZACIÓN DEL GEN <i>tra2</i> DE <i>S. coprophila</i></b>	<b>41</b>
3.2.3.1 Amplificación mediante PCR de un fragmento de ADN genómico	41
3.2.3.2 RACE ("Rapid Amplification of cDNA Ends")	41
3.2.3.3 <i>Genome Walker</i> (Paseo Cromosómico)	42
3.2.3.4 RT-PCR con cebadores específicos del gen <i>Sctra2</i>	42
3.2.3.5 Expresión del gen <i>tra2</i> de <i>S. coprophila</i> durante el desarrollo mediante RT-PCR	43
<b>3.2.4 MOSCAS TRANSGÉNICAS De <i>Drosophila melanogaster</i></b>	<b>44</b>
3.2.4.1 Construcción del elemento transgénico <i>TDSO</i>	44
3.2.4.2 Generación de moscas transgénicas	45
3.2.4.3 Localización cromosómica de los transgenes	45
3.2.4.4 Análisis de la expresión de los transgenes	47
3.2.4.6 Efecto del transgén <i>TDSO</i> sobre el procesamiento del pre-ARNm del gen <i>doublesex</i> de <i>Drosophila melanogaster</i> .	47
<b>4. RESULTADOS</b>	<b>53</b>
<b>4.1 AISLAMIENTO, ORGANIZACIÓN MOLECULAR Y TRANSCRITOS DEL GEN <i>tra2</i> DE <i>S. ocellaris</i></b>	<b>54</b>
<b>4.2 AISLAMIENTO, ORGANIZACIÓN MOLECULAR Y TRANSCRITOS DEL GEN <i>tra2</i> DE <i>S. coprophila</i></b>	<b>59</b>
<b>4.3 COMPARACIÓN DE LA ORGANIZACIÓN MOLECULAR DEL GEN <i>tra2</i> DE <i>Sciara ocellaris</i> Y <i>Sciara coprophila</i> CON EL GEN <i>tra2</i> DE OTROS INSECTOS.</b>	<b>63</b>
<b>4.4 LA PROTEÍNA TRA2 DE <i>S. ocellaris</i> Y <i>S. coprophila</i> Y SU COMPARACIÓN CON LAS PROTEÍNAS TRA2 DE OTROS INSECTOS.</b>	<b>65</b>
<b>4.5 PATRÓN DE EXPRESIÓN DEL GEN <i>tra2</i> DE <i>S. ocellaris</i> Y <i>S. coprophila</i></b>	<b>69</b>
<b>4.6 EL TRANSCRITO <i>Sotra2-130</i> ESTÁ PRESENTE TANTO EN MACHOS COMO EN HEMBRAS.</b>	<b>71</b>
<b>5. DISCUSIÓN</b>	<b>78</b>
<b>5.1 ANÁLISIS COMPARATIVO DE LAS PROTEÍNAS TRA2 DE INSECTOS</b>	<b>79</b>
<b>5.2 CONSERVACIÓN EVOLUTIVA DE LA FUNCIÓN DEL GEN <i>tra2</i> DE <i>Sciara</i> EN LA DETERMINACIÓN SEXUAL DE <i>Drosophila</i></b>	<b>81</b>
<b>5.3 FUNCIÓN Y EVOLUCIÓN DEL GEN <i>tra2</i></b>	<b>83</b>
<b>6. CONCLUSIONES</b>	<b>87</b>
<b>7. BIBLIOGRAFÍA</b>	<b>89</b>
<b>8. ANEXOS</b>	<b>102</b>

**µg:** microgramo

**µl:** microlitro

**µM:** micromolar

**aa:** aminoácido

**ADN:** ácido desoxirribonucleico

**ADNc:** ácido desoxirribonucleico complementario

**ARN:** ácido ribonucleico

**ARNm:** ácido ribonucleico mensajero

**°C:** grados centígrados

**CyO:** alelo mutante del gen *Curly-o*

**dNTP:** desoxinucleótido trifosfato

**dsx:** *doublesex*

**DsxRE:** elemento repetido de *doublesex* (*doublesex repeat element*)

**EDTA:** ácido etilendiamintetraacético

**fru:** *fruitless*

**g:** gramo

**Gal:** galactosidasa

**HS:** choque térmico (*Heat shock*)

**IPTG:** isopropil β-D-1-tiogalactopiranosido

**ix:** *intersex*

**l:** litro

**LB:** medio de Luria-Bertani

**M:** molar

**mg:** miligramo

**MgCl<sub>2</sub>:** cloruro de magnesio

**ml:** mililitro

**mM:** milimolar

**ng:** nanogramo

**ORF:** marco abierto de lectura (*open reading frame*)

**pb:** pares de bases

**PCR:** reacción en cadena de la polimerasa (*polymerase chain reaction*)

**poli-A:** ácido poliadenílico

**poli-U:** ácido poliuracílico

**PRE:** potenciador rico en purinas (*purine rich enhancer*)



**RACE:** amplificación rápida de los extremos de ADNc (*rapid amplification cDNA ends*)

**RT:** retrotranscripción

**Sb:** alelo mutante del gen *Stubble*

**SDS:** dodecil sulfato sódico (*sodium dodecyl sulfate*)

**Sxl:** *Sex-lethal*

**TBE:** tampón tris bórico EDTA

**TE:** tampón tris EDTA

**Tris:** Tris(hidroximetil)amino-metano

**tra:** *transformer*

**tra2:** *transformer-2*

**tra2<sup>B</sup>:** alelo mutante del gen *tra2*

**U:** unidad

**UTR:** región no traducida (*untranslated region*)

**w:** alelo mutante del gen *white*

**w<sup>+</sup>:** alelo silvestre del gen *white*

**y:** alelo mutante del gen *yellow*

## **1. INTRODUCCIÓN**

La perpetuación de las especies a través de la reproducción sexual es la regla general dentro del Reino Animal. Las hembras y los machos son diferentes desde el punto de vista morfológico, fisiológico y de conducta. Este dimorfismo sexual es el resultado de la integración de dos procesos: determinación sexual y diferenciación sexual. La determinación sexual es el proceso que hace que el embrión siga un desarrollo de macho o de hembra. Los genes responsables se denominan genes de determinación sexual. La diferenciación sexual hace referencia a la expresión de los genes de citodiferenciación, controlados por los genes de determinación sexual, cuya expresión da lugar a la formación de las estructuras sexuales dimórficas que caracterizan a hembras y machos adultos.

## **1.1 MECANISMOS DE DETERMINACIÓN SEXUAL EN INSECTOS**

En el Reino Animal existen distintos mecanismos de determinación sexual, estando todos ellos representados en los insectos. No obstante, estos mecanismos pueden englobarse en tres tipos básicos en función de cual sea la señal primaria que determine el sexo: constitución cromosómica o genética del cigoto, efecto materno y condiciones ambientales.

La determinación sexual puede estar definida por la constitución cromosómica del cigoto. Este mecanismo implica diferencias cromosómicas, siendo un sexo homomórfico y el otro sexo heteromórfico, para los cromosomas sexuales. Así, hay insectos en los que las hembras son el sexo homomórfico (XX) y los machos son el sexo heteromórfico (XY). El ejemplo más representativo de esta situación tiene lugar en *Drosophila melanogaster* (la mosca del vinagre), donde la determinación sexual está basada en la razón entre el número de cromosomas X y el número de juegos haploides de autosomas, siendo las hembras 2X;2A y los machos X;2A (X hace referencia al cromosoma X y A hace referencia al complemento haploide de autosomas). En otras especies, el sexo heteromórfico es portador de un factor determinante de macho que puede estar localizado en el cromosoma Y, como es el caso de los trefritidos *Ceratitis capitata* (la mosca mediterránea de la fruta), *Bactrocera oleae* (la mosca del olivo) y *Anastrepha obliqua* (la mosca de la fruta) o de *Musca domestica* (la mosca doméstica), si bien en esta última, el factor de masculinidad puede

encontrarse en un autosoma. Sin embargo, existen también insectos, como los lepidópteros, en los que ocurre lo contrario, siendo los machos (ZZ) el sexo homomórfico y las hembras (ZW) el sexo heteromórfico. En otros casos, las diferencias cromosómicas son debidas a la existencia de un sistema haplo-diploide como es el caso de *Apis mellifera* (la abeja), siendo las hembras diploides y los machos haploides.

En la mayoría de las especies, entre las que se incluyen las anteriormente mencionadas, la constitución cromosómica del cigoto es consecuencia directa de la constitución cromosómica de los gametos (Bull, 1983). Sin embargo, en otras especies, las diferencias cromosómicas responsables de determinar el sexo son consecuencia del comportamiento especializado de los cromosomas sexuales en los primeros estadios del desarrollo embrionario. En las especies de dípteros de *Sciara* (mosquito del hongo), todos los cigotos comienzan con una constitución cromosómica 3X;2A (donde dos cromosomas X provienen del padre y un cromosoma X de la madre) produciéndose en éstos, una eliminación diferencial de los cromosomas sexuales. La pérdida de un cromosoma X paterno origina un cigoto 2X;2A que dará lugar a una hembra, mientras que la pérdida de los dos cromosomas X paternos origina un cigoto X0;2A que dará lugar a un macho. En los cóccidos (la cochinilla), tanto la eliminación diferencial de los cromosomas paternos, como la heterocromatinización de todos los mismos, originan cigotos diploides o haploides (estructurales o funcionales) que se desarrollan como hembras o machos, respectivamente.

A nivel genético, el sexo puede estar controlado por un único locus o por varios loci (Bull, 1983). En el caso de la abeja, un único locus con varios alelos, llamado *Complementary Sex Determination* (CSD), determina el sexo del cigoto. Las hembras son siempre heterocigotas para este locus mientras que los machos fértiles son hemicingóticos (haploides) y los machos estériles son homocigóticos (diploides) (Beye *et al.*, 2003).

Un ejemplo de determinación sexual debida a un efecto materno ocurre en *Chrysomya rufifacies* (la mosca azul o mostarda), donde el sexo del cigoto es determinado, exclusivamente, por el genotipo de la madre. En esta especie existen dos tipos de hembras. Las hembras ginogénicas son heterocigóticas para el gen *F*, que codifica un factor materno depositado en los oocitos durante la oogénesis y que impone el desarrollo de hembra a los cigotos derivados de este oocito. Las hembras androgénicas, al igual que los machos, son

homocigóticas para el alelo recesivo  $f$ , que no produce factor materno, por lo que sólo producen machos en su descendencia.

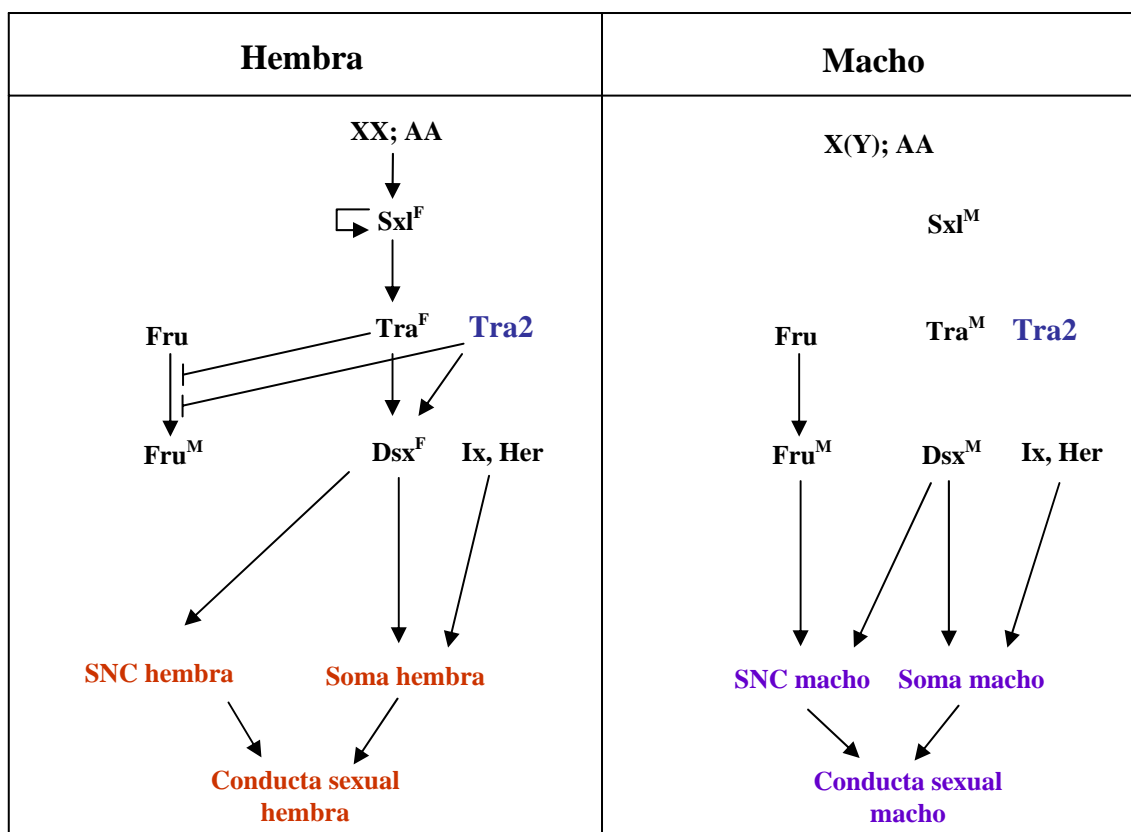
Finalmente, los factores ambientales también pueden ser los responsables de determinar el sexo. Es conocido el efecto de la temperatura en algunas especies de ciáridos, donde la razón sexual (machos *versus* hembras) puede desviarse del valor 1 que toma a nivel poblacional.

Todos estos mecanismos ocurren en los insectos. En *Sciara* los tres tipos de mecanismos se dan de forma concatenada, por lo que este organismo constituye un modelo experimental clave para el estudio integrado de los distintos mecanismos de determinación sexual que han ido apareciendo a lo largo de la evolución.

## 1.2 DETERMINACIÓN SEXUAL EN *Drosophila melanogaster*

El sistema de determinación sexual de referencia es el que presenta *D. melanogaster* ya que ha sido minuciosamente analizado. Las relaciones epistáticas existentes entre los genes de la determinación sexual revelan que se encuentran organizados jerárquicamente (Figura 1), de manera que el producto de un gen controla el procesamiento específico de sexo del pre-ARNm del gen que se encuentra por debajo en la cascada (Sánchez *et al.*, 2005). El gen *Sex-lethal* (*Sxl*) ocupa la posición superior en la cascada; su producto controla el procesamiento de su propio pre-ARNm, así como el procesamiento del pre-ARNm del gen *transformer* (*tra*). La proteína Tra junto con el producto del gen constitutivo *transformer-2* (*tra2*) controlan el procesamiento específico de sexo del pre-ARNm del último gen en la cascada, *doublesex* (*dsx*), que se transcribe en ambos sexos dando lugar a proteínas diferentes, DsxM y DsxF.

La señal primaria que determina el sexo es la señal X/A (razón entre el número de cromosomas X y el número de juegos haploides de autosomas). En individuos 2X;2A esta señal X/A tiene un valor de 1, lo que determina el desarrollo de hembra, mientras que en individuos 1X;2A la señal X/A tiene un valor de 0.5, lo que determina el desarrollo de macho (Cline, 1993). Recientemente, se ha propuesto que es el número de cromosomas X y la relación X/A la señal primaria que determina el sexo (Erickson y Quintero, 2007).



**Figura 1. Cascada genética de la determinación sexual en *D. melanogaster*.** Sxl<sup>F</sup> y Sxl<sup>M</sup> indican proteína Sxl funcional y no funcional, respectivamente. Tra<sup>F</sup> y Tra<sup>M</sup> indican proteína Tra funcional y no funcional, respectivamente. Dsx<sup>F</sup> y Dsx<sup>M</sup> indican proteína Dsx funcional de hembra y macho, respectivamente. Fru<sup>M</sup> indica proteína Fru funcional específica de macho. En los machos, el valor de la señal X/A, conlleva a que se produzcan por defecto las proteínas Sxl<sup>M</sup>, Tra<sup>M</sup>, Dsx<sup>M</sup> y Fru<sup>M</sup>. En hembras, la señal X/A tiene un valor de 1 y se producen proteínas Sxl<sup>F</sup>, Tra<sup>F</sup> y Dsx<sup>F</sup>. Las proteínas Tra2, Ix y Her son producidas en ambos sexos. SNC hace referencia al sistema nervioso central (Sánchez, 2008).

El gen *Sex-lethal* (*Sxl*) tiene dos promotores, un promotor temprano y un promotor tardío (Salz *et al.*, 1989). En respuesta a la señal X/A el promotor temprano del gen se activa, sólo en los cigotos 2X;2A, lo que da lugar a la formación de proteína Sxl temprana sólo en las hembras. Una vez se llega al estadio de blastodermo, la señal X/A ya no se requiere y la actividad de *Sxl* queda fijada (Sánchez y Nöthiger, 1983; Bachiller y Sánchez, 1991) gracias a su capacidad de autorregulación. Esta capacidad de autorregulación positiva, por medio de la cual la proteína Sxl participa en el procesamiento de su propio transcrito primario, proporciona la memoria celular para el desarrollo de hembra (Cline, 1984).

Más tarde en el desarrollo, después del blastodermo, el promotor tardío de este gen se activa en ambos sexos, produciéndose transcrito tardío de *Sxl* durante el resto del desarrollo y la vida adulta. Los transcritos tardíos de los machos son

idénticos a los de las hembras, excepto por la presencia en ellos de un exón (L3) adicional que contiene un codón de parada de la traducción, lo que produce una proteína Sxl truncada y no funcional. En las hembras, se produce un procesamiento alternativo por la unión de la proteína Sxl temprana a las secuencias poli-U localizadas en los intrones 2 y 3 que flanquean al exón específico de macho (L3), produciéndose un ARNm que no contiene dicho exón y que, por lo tanto, da lugar a una proteína Sxl tardía funcional (Bell *et al.*, 1988). La proteína Sxl va a regular el procesamiento alternativo específico de sexo del pre-ARNm del gen *transformer* (*tra*).

El gen *tra* se transcribe en machos y hembras, pero su transcrito primario sigue un procesamiento alternativo específico de sexo de manera que sólo en las hembras se produce proteína Tra. El pre-ARNm del gen *tra* tiene dos sitios 3' de procesamiento alternativo en el exón 2, uno específico de hembra (distal) y otro no específico de sexo (proximal). Cuando éste último es utilizado, se genera un transcrito que contiene un codón de parada de la traducción dentro del marco abierto de lectura, generándose, por tanto, una proteína Tra truncada y no funcional. En las hembras, aproximadamente la mitad del pre-ARNm de *tra* es procesado de forma diferente, debido a la presencia de la proteína Sxl, que compite con el factor de procesamiento U2AF (U2 auxiliary factor) por unirse al tracto de polipirimidinas del sitio 3' proximal. El factor U2AF, al no poder unirse a este sitio 3' proximal, se va a unir al tracto de polipirimidinas del sitio 3' distal, específico de hembra, por el que presenta una menor afinidad. De esta forma, el codón de parada de la traducción no es introducido en el ARNm y se origina una proteína Tra completa y funcional (Boggs *et al.*, 1987; Belote *et al.*, 1989; Valcárcel *et al.*, 1993). La proteína Tra, junto con el producto del gen constitutivo *transformer-2* (*tra2*), controla el procesamiento específico de sexo del pre-ARNm del gen *doublesex* (*dsx*).

El gen *dsx* ocupa la última posición de la jerarquía genética que controla la determinación sexual. Se transcribe en ambos sexos dando lugar a transcritos diferentes por un procesamiento alternativo específico de sexo. En las hembras, la unión del complejo Tra-Tra2 al pre-ARNm determina la incorporación de un exón específico de hembra al ARNm. Por medio de este procesamiento tipo hembra, se origina la proteína DsxF, responsable del desarrollo sexual femenino. En los machos, la ausencia de la proteína Tra hace que no se forme el complejo Tra-Tra2 y el procesamiento tiene lugar por defecto, incorporándose los exones específicos de macho al ARNm, originándose la

proteína DsxM que determina el desarrollo sexual de macho. Como consecuencia, se obtienen dos proteínas, ambas funcionales: DsxM, en los machos, y DsxF, en las hembras, con una función antagónica en la regulación transcripcional de los genes responsables de la diferenciación sexual terminal.

El complejo Tra-Tra2, controla también el procesamiento alternativo específico de sexo del pre-ARNm del gen *fruitless* (*fru*). Este gen, al igual que *dsx*, está involucrado en el desarrollo sexual masculino del sistema nervioso central (Rideout *et al.*, 2007) requerido para el cortejo en los machos (Shirangi *et al.*, 2006). El gen *fru* es un gen complejo que presenta cuatro promotores. Se transcribe en ambos sexos generando distintos tipos de transcritos por procesamiento alternativo de su transcrito primario. El promotor P1 únicamente funciona en un grupo reducido de neuronas del sistema nervioso central, en varias regiones del cerebro, y en el ganglio ventral (Billeter *et al.*, 2002; 2006). El procesamiento alternativo del pre-ARNm transcrito a partir de este promotor está regulado por el complejo Tra-Tra2. En las hembras, la unión de este complejo al pre-ARNm determina la incorporación de un exón específico de hembra que contiene codones de parada de la traducción, generándose una proteína Fru no funcional. En los machos, donde la proteína Tra no está presente, el exón específico de hembra no se incorpora, produciéndose, entonces, una proteína FruM funcional (Ryner *et al.*, 1996; Heinrichs *et al.*, 1998; Goodwin *et al.*, 2000). Los promotores P2-P4 funcionan en tejidos neuronales y no neuronales en estadios embrionarios y codifican proteína funcional en ambos sexos.

El gen *hermaphrodite* (*her*) tiene una dualidad de función. Su expresión materna es necesaria para la activación temprana de *Sxl*, mientras que su expresión cigótica es necesaria para la diferenciación terminal de las hembras y algunos aspectos de la diferenciación terminal de los machos (Pultz y Baker, 1995; Li y Baker, 1998).

El gen *intersex* (*ix*) se transcribe en ambos sexos dando lugar a un pre-ARNm que no sigue un procesamiento específico de sexo, por lo que la proteína Ix está presente en machos y en hembras. Se ha demostrado la interacción de esta proteína con DsxF, pero no con DsxM, formándose un complejo con capacidad de unión al ADN que estaría implicado en la activación de los genes de diferenciación sexual de hembra (Chase y Baker, 1995; Waterbury *et al.*, 1999; Garrett-Engle *et al.*, 2002).



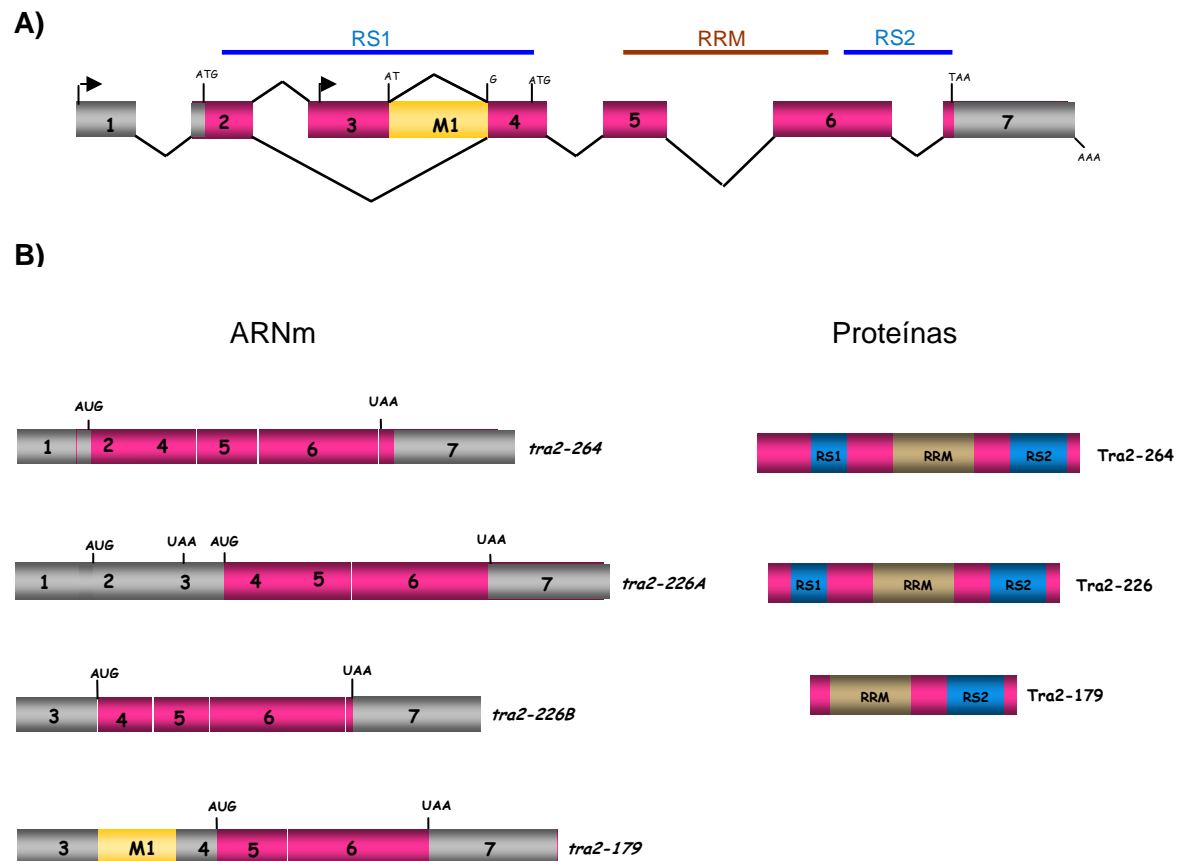
### 1.3 EL GEN *tra 2* DE *Drosophila melanogaster*

El gen *tra2* de *Drosophila melanogaster* está compuesto por siete exones (Figura 2A), se transcribe en ambos sexos y da lugar a cuatro ARNm (*tra2-264*, *tra2-226A*, *tra2-226B* y *tra2-179*) distintos debido al uso de diferentes promotores y a un procesamiento alternativo (Figura 2B). Los transcritos, *tra2-264* y *tra2-226A* se expresan en el soma de machos y hembras, y en la línea germinal de las hembras de manera constitutiva, a lo largo de todo el desarrollo y durante la vida adulta (Amrein *et al.*, 1988; Goralski *et al.*, 1989; Mattox y Baker, 1991). Los transcritos, *tra2-226B* y *tra2-179*, son específicos de la línea germinal de los machos.

El transcrito *tra2-264* constituye la forma más abundante, está formado por los exones E1, E2, E4, E5, E6 y E7 y codifica una proteína de 264 aminoácidos (Tra2-264) que comienza en el codón AUG situado en el exón E2.

El transcrito *tra2-226A* de la línea somática, está constituido por los exones E1, E2, E3, E4, E5, E6 y E7. La inclusión del exón 3 en el ARN mensajero maduro introduce codones de parada de la traducción en fase con el codón AUG del exón E2. Sin embargo, el procesamiento del exón 3 al exón 4 crea un nuevo AUG que es utilizado para el inicio de la traducción, dando lugar a un polipéptido de 226 aminoácidos (Tra2-226), con el mismo marco abierto de lectura que Tra2-264, pero que carece de parte del extremo amino terminal.

Los dos transcritos específicos de la línea germinal de los machos comienzan su transcripción a partir de un promotor localizado en el exón 3. El transcrito *tra2-226B*, está formado por los exones E3, E4, E5, E6 y E7, codificando una proteína de 226 aminoácidos (Tra2-226) idéntica a la proteína de la línea somática codificada por el ARNm *tra2-226A*. Por último, el transcrito *tra2-179* es igual que el anterior, salvo por la presencia en este del intrón M1, el cual introduce codones de parada de la traducción, dando lugar a una proteína truncada de 179 aminoácidos (Tra2-179).



**Figura 2. Estructura y patrón de expresión del gen *tra2* de *D. melanogaster*.** Los rectángulos coloreados en rosa representan los exones. El rectángulo amarillo representa el intrón M1 que se retiene en la línea germinal. Las regiones 5'UTR y 3'UTR se representan en un tono más claro. Se indican los codones de inicio y parada de la traducción, los sitios de poliadenilación (AAA) y los dominios RS1, RS2 (líneas azules) y RRM (línea marrón). Las líneas de color negro representan el procesamiento alternativo. Las flechas señalan el inicio de la transcripción. **(A)** Organización molecular del gen. **(B)** Esquema de los ARNm y sus proteínas.

### 1.3.1 Las proteínas Tra2

Las tres isoformas de la proteína Tra2 de *D. melanogaster* (Tra2-264, Tra2-226 y Tra2-179) (Figura 2B) pertenecen a la familia de proteínas SR (*Splicing Regulator*), cuyos miembros pueden estar implicados en la regulación del procesamiento, transporte de ARN mensajero y traducción (Shepard y Hertel, 2009). Estas proteínas se caracterizan por su capacidad de interaccionar simultáneamente con el ARN y otras proteínas a través de los dominios RRM (RNA recognition motif) y RS (regiones ricas en serina y arginina), respectivamente. Las tres isoformas se diferencian en su tamaño. El extremo amino-terminal de la proteína Tra2-226 que precede al dominio RS1 es más

corto que el de la proteína Tra2-264. El extremo amino-terminal de la proteína Tra2-179 es aún más corto, faltándole, además, el dominio RS1 (ver Figura 2B) (Amrein et al., 1994; Mattox et al., 1996 y Wu y Maniatis, 1993).

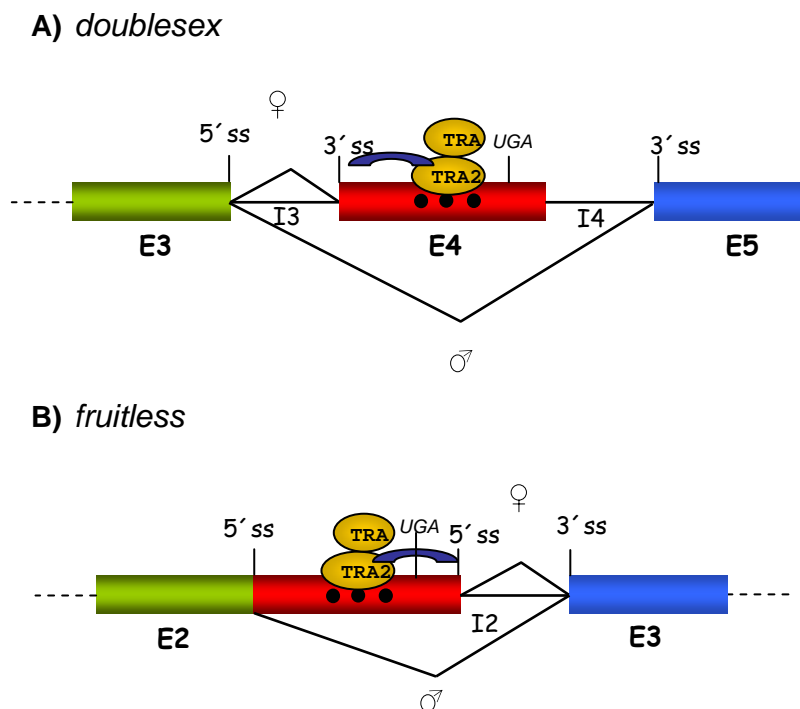
Estas proteínas se unen a las secuencias ESE (exonic splicing enhancer) presentes en los exones de los genes que sufren un procesamiento. Estas secuencias ESE son normalmente reconocidas por al menos un miembro de la familia de proteínas RS para reclutar así la maquinaria de procesamiento al intrón adyacente (Black, 2003; Blencowe, 2000; Gravely, 2000; Gravely *et al.*, 2001).

### 1.3.2 Función

En *Drosophila*, la proteína Tra2 muestra una dualidad de función. Se comporta como un activador o un inhibidor del procesamiento. En los tejidos somáticos, las isoformas Tra2-264 y Tra2-226 actúan de manera redundante (Amrein *et al.*, 1994; Mattox *et al.*, 1996) y, junto con la proteína Tra, regulan el procesamiento alternativo específico de sexo de los transcritos primarios de los genes *doublesex* (Hoshijima *et al.*, 1991) y *fruitless* (Ryner *et al.*, 1996) conduciendo a la expresión específica de sexo de los factores de transcripción codificados por estos genes que regulan la identidad sexual de los diferentes tejidos somáticos (Heinrichs *et al.*, 1998; Tian y Maniatis 1992).

En las hembras, la proteína Tra2 forma un complejo con la proteína Tra. Este complejo se une, en la primera mitad del exón 4 del pre-ARNm del gen *dsx*, al elemento DsxRE (*doublesex* repeat element) y al elemento PRE (purine-rich element), reclutando al factor de procesamiento U2AF y a otros componentes de la maquinaria general de procesamiento (Figura 3A). Como consecuencia, el sitio 3' de procesamiento del intrón 3, que es un sitio débil porque su secuencia se aleja de la consenso, es reconocido y el exón 4 es incorporado al ARNm. Esto conduce a la formación de un mensajero que codifica la proteína DsxF, responsable de la diferenciación de hembra. En los machos, debido a la ausencia de Tra, el procesamiento tiene lugar, por defecto, usándose el sitio 3' del intrón 4, no incorporándose el exón 4 al ARNm y sí los exones 5 y el 6, dando lugar a la proteína DsxM, que determina el desarrollo sexual de macho (Tian y Maniatis, 1993).

En el sistema nervioso central, las proteínas Tra2 y Tra regulan el procesamiento alternativo específico de sexo del pre-ARNm del gen *fruitless* (*fru*) producido a partir del promotor P1 (Ver Introducción). En las hembras, el complejo Tra-Tra2 se une a las secuencias ESE localizadas en el exón 2. (Figura 3B) y, como consecuencia, el sitio 5' de procesamiento del intrón 2, que es un sitio débil porque su secuencia se aleja de la consenso, es reconocido y el exón específico de hembra es incorporado al ARNm. Este exón contiene codones de parada de la traducción, generándose una proteína Fru no funcional. En los machos, donde la proteína Tra no está presente, el procesamiento tiene lugar por defecto, y el exón específico de hembra no se incorpora al ARNm. Se produce, entonces, una proteína FruM funcional (Ryner *et al.*, 1996; Heinrichs *et al.*, 1998; Goodwin *et al.*, 2000).



**Figura 3. Regulación del procesamiento alternativo específico de sexo del pre-ARNm de los genes *doublesex* (A) y *fruitless* (B) de *D. melanogaster*.** Los rectángulos coloreados representan los exones. En verde el exón común a ambos sexos, en rojo el exón específico de hembra y en azul el exón específico de macho. Las líneas negras indican el procesamiento alternativo específico de hembra o de macho. Los puntos negros indican los sitios de unión del complejo Tra-Tra2. Se indican los codones de parada de la traducción y los sitios 5' y 3' de procesamiento (ss).

En la línea germinal de los machos de *D. melanogaster*, se expresan las proteínas Tra2-226 (requerida para la fertilidad de los machos) y la proteína truncada Tra2-179, presumiblemente no funcional (Amrein *et al.*, 1990; Mattox *et al.*, 1990). En este caso, y de manera independiente de Tra, la proteína Tra2 regula el procesamiento alternativo de su propio transcrito primario (Hazelrigg y Tu 1994; Madigan *et al.*, 1996; McGuffin *et al.*, 1998). En ausencia de la proteína Tra2, las células germinales de los machos parecen iniciar una espermatogénesis normal pero, al final, forman espermátidas con defectos en la elongación nuclear (Belote y Baker, 1993). Como resultado, estos machos mutantes son infértiles.

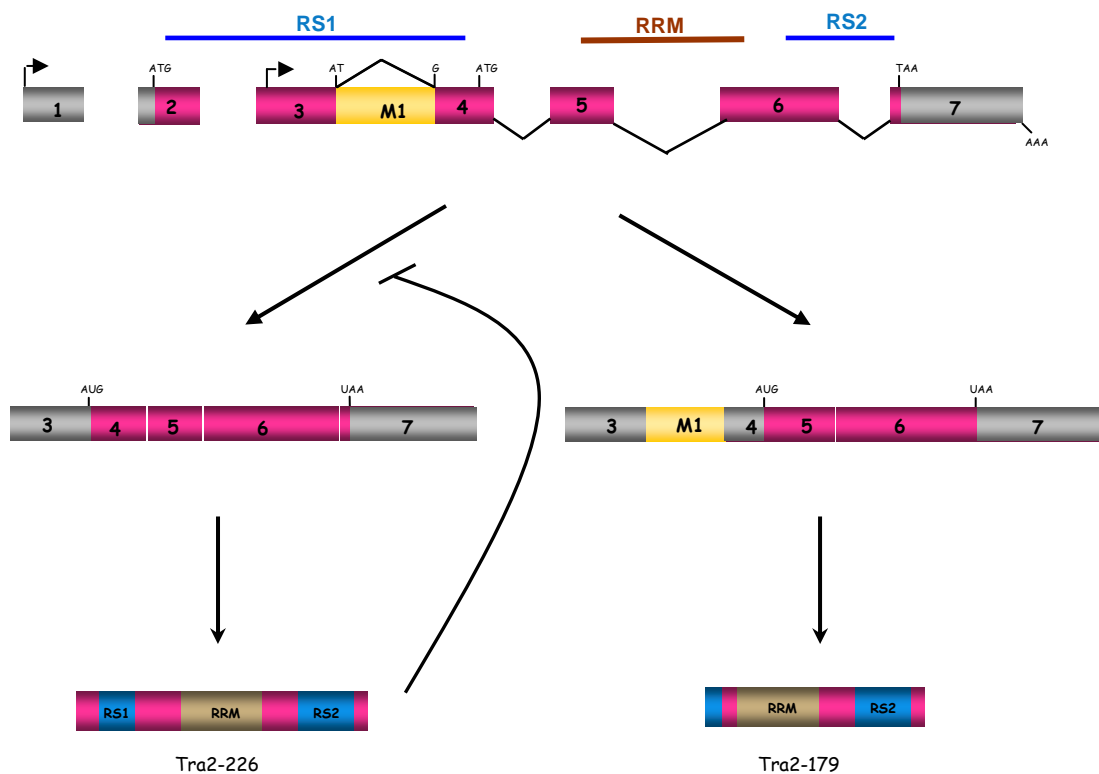
Los niveles relativos de los ARNm que codifican estas isoformas están controlados por un mecanismo de autorregulación en el que la isoforma Tra2-226 previene el procesamiento del intrón M1 de su propio transcrito primario (Figura 4) (Mattox y Baker 1991; Mattox *et al.*, 1996). Este intrón interrumpe el codón de inicio de la traducción de la propia proteína Tra2-226 actuando, de esta manera, como un regulador negativo de su propia expresión. Cuando el procesamiento del intrón M1 es reprimido, el ARN mensajero de *tra2* que se acumula, codifica una isoforma truncada de la proteína (Tra2-179) que carece del dominio RS1 de la región amino terminal. Esta retención del intrón M1 es el mecanismo por el que la proteína Tra2 funcional limita su propia síntesis ya que un exceso de proteína provoca esterilidad en los machos (McGuffin *et al.*, 1998).

El procesamiento del intrón M1 es mucho menos eficiente que el de los otros intrones de Tra2 debido a la existencia de un sitio 3' de procesamiento que se aleja de la secuencia consenso (Chandler *et al.*, 2001), haciendo que los transcritos que contienen el intrón M1 en la línea germinal de los machos supongan un 60-70% del total (Amrein *et al.*, 1990; Mattox *et al.*, 1990).

La proteína Tra2-226 evita el procesamiento del intrón M1 de su transcrito primario, uniéndose a una región específica, localizada en este intrón, denominada ISS (intronic silencer splicing). Cinco repeticiones de la secuencia CAAGR en esta región son necesarias para la unión y represión del procesamiento por Tra2 (Chandler *et al.*, 2003; Mattox y Baker 1991; Qi *et al.*, 2007). Estas secuencias difieren de las secuencias repetidas de los sitios de unión encontrados en los ESE de los genes *dsx* y *fru*. El ISS también contiene una secuencia consenso de unión de la proteína RBP1. Sin, embargo, al

contrario de lo que ocurre en *dsx*, donde Tra2 y RBP1 se unen a las secuencias ESE de manera cooperativa, en este caso Tra2 y RBP1 se unen al ISS independientemente a través de distintas secuencias.

Ensayos *in vitro* sugieren que esta interacción entre Tra2, RBP1 y las ISS no es suficiente para la inhibición del procesamiento del intrón M1 sino que se requiere la presencia de extractos nucleares lo que sugiere que otros factores adicionales deben estar también implicados. Este factor no puede ser la proteína Tra ya que esta no se produce en los machos de *Drosophila* (Qi *et al.*, 2007).



**Figura 4. Regulación del procesamiento del pre-ARNm del gen *tra2* en la línea germinal de los machos de *D. melanogaster*.** Los rectángulos coloreados en rosa representan los exones. El rectángulo amarillo representa el intrón M1 que se retiene en la línea germinal. Las regiones 5'UTR y 3'UTR se representan en un tono más claro. Se indican los codones de inicio y parada de la traducción, los sitios de poliadenilación (AAA) y los dominios RS1, RS2 (líneas azules) y RRM (línea marrón). Las líneas de color negro representan el procesamiento alternativo. La flecha negra señala el inicio de la transcripción.

## 1.4 GENES DE LA DETERMINACIÓN SEXUAL EN OTROS INSECTOS

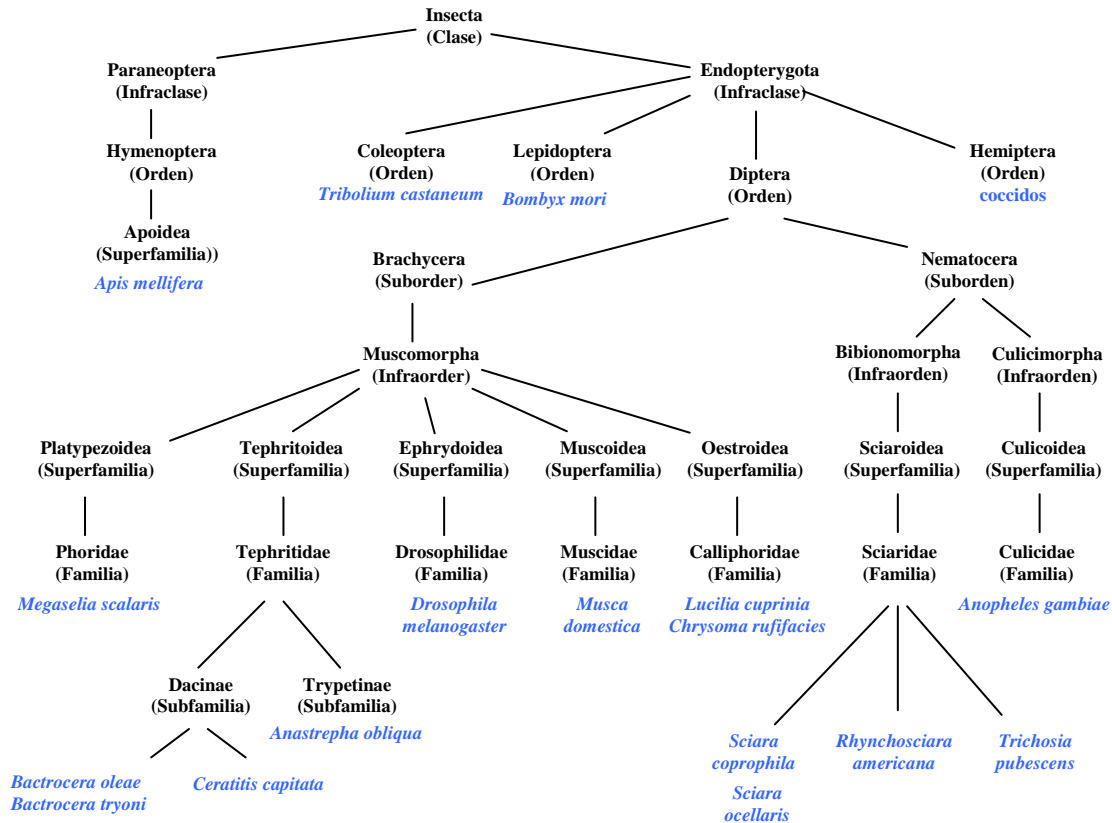
El estudio de la evolución de los mecanismos que controlan la determinación sexual en los insectos ha sido abordado mediante el aislamiento en otros insectos (Figura 5) de los homólogos de los genes que controlan este proceso en *D. melanogaster*. La idea subyacente es determinar cómo ha podido cambiar la cascada de los genes de la determinación sexual a lo largo de la evolución, desde los insectos más primitivos a los insectos más evolucionados como son los drosofilidos. A continuación se resume el estado actual de este análisis.

### 1.4.1 El gen *sex-lethal*

El gen *sxl* ha sido caracterizado en otras especies de *Drosophila*: *D. virilis* (Bopp *et al.*, 1996) y *D. pseudoobscura* (Penalva *et al.*, 1996). Como en *D. melanogaster*, la regulación de *sxl* ocurre por un procesamiento alternativo específico de sexo: el ARNm en machos tiene un exón adicional que contiene codones de parada de la traducción. En *D. virilis*, *sxl* presenta, a continuación del último codón de parada de la traducción de este exón, otro marco abierto de lectura, generando una proteína Sxl que es idéntica a la de las hembras excepto por los primeros 25 aminoácidos de la región amino terminal. La proteína Sxl de macho es acumulada fundamentalmente en el ectodermo del embrión, lo que sugeriría un papel en el desarrollo del sistema nervioso central (Bopp *et al.*, 1996). Esta misma proteína también ha sido detectada en otras especies (*D. americana*, *D. flavomontana* y *D. borealis*) (Bopp *et al.*, 1996).

Fuera del género *Drosophila*, *Sxl* ha sido caracterizado en los dípteros *Chrysomya rufifacies* (Müller-Holtkamp, 1995), *Megaselia scalaris* (mosca fórico) (Sievert *et al.*, 1997, 2000), *Musca domestica* (Meise *et al.*, 1998) y los tefritidos *Ceratitis capitata* (Saccone *et al.*, 1998) y *Bactrocera oleae* (Lagos *et al.*, 2005), pertenecientes al suborden Brachycera, y en *Sciara ocellaris* (Ruiz *et al.*, 2003), *Sciara coprophila*, *Rynchosciara americana* y *Trichosia pubescens* (Serna *et al.*, 2004), pertenecientes al suborden Nematocera. *Sxl* también ha sido caracterizado en el lepidóptero *Bombyx mori* (gusano de la seda) (Niimi *et al.*, 2006). El gen *Sxl* de estas especies no muestra una regulación específica de sexo, produciéndose la misma proteína Sxl en machos y en hembras. Así, *Sxl* no parece jugar el papel clave en el control de la determinación sexual que

tiene en los drosofilidos, lo que sugiere que *sxl* ha adquirido esta función durante la evolución del linaje de *Drosophila*.



**Figura 5.** Clasificación de los insectos mencionados en esta tesis. En color azul se muestran las especies mencionadas. (Modificado de Sánchez, 2008).

#### 1.4.2 El gen *transformer*

El gen *tra* ha sido caracterizado en las siguientes especies del género *Drosophila*: *D. simulans*, *D. mauritiana*, *D. sechellia*, *D. erecta* (O'Neil y Belote, 1992; Kulathinal *et al.*, 2003), *D. hydei* y *D. virilis*. Su comparación con el gen *tra* de *D. melanogaster* revela un elevado grado de divergencia. El gen *tra* de *D. virilis* provee parcialmente la función *tra*<sup>+</sup> en moscas de *D. melanogaster* mutantes para este gen (O'Neil y Belote, 1992).

Fuera de los drosofilidos, el gen *tra* ha sido caracterizado, en los tefrítidos *C. capitata* (Pane *et al.*, 2002; 2005), *B. oleae* (Lagos *et al.*, 2007) y en doce



especies de *Anastrepha*: *A. obliqua*, *A. serpentina*, *A. striata*, *A. bistrigata*, *A. grandis*, *A. amita*, *A. sororcula*, *A. ludens*, *A. sp1. aff. fraterculus*, *A. sp2. aff. fraterculus*, *A. sp3. aff. fraterculus* y *A. sp4. aff. Fraterculus* (Ruiz *et al.*, 2007a). En todos estos tefrítidos, el gen *tra* juega un papel regulador clave, actuando como la memoria celular para la determinación sexual mediante un mecanismo de autorregulación positiva en el que la propia proteína Tra regula el procesamiento específico de hembra de su propio transcrito primario (Pane *et al.*, 2002). De este modo, sólo en las hembras se produce proteína Tra funcional, mientras que en los machos se produce proteína Tra truncada no funcional. Además, el gen *tra* de *C. capitata* (Pane *et al.*, 2005) y de *Anastrepha obliqua* (Ruiz y Sánchez, 2010) provee parcialmente la función *tra* en moscas de *D. melanogaster* mutantes para este gen.

El gen *tra* también es esencial para el desarrollo de hembra en *Lucilia cuprina* (Diptera, Calliphoridae) (Concha y Scout, 2009) y *M. domestica* (Diptera, Muscidae) (Hediger *et al.*, 2010). Este gen, también muestra autorregulación positiva, de modo que sólo en las hembras se produce una proteína Tra funcional.

#### **1.4.3 El gen *transformer-2***

Además de en *D. melanogaster*, el gen *tra2* se ha caracterizado en *D. virilis* (Chandler *et al.*, 1997). En esta especie, *tra2* codifica varias isoformas análogas a las de la proteína Tra2 de *D. melanogaster*. Además, en moscas de *D. melanogaster* mutantes para *tra2*, el gen *tra2* de *D. virilis* es capaz de reemplazar la función del gen endógeno, tanto para la diferenciación sexual de hembra, regulando el procesamiento alternativo del gen *dsx* endógeno, como para la espermatogénesis, regulando el procesamiento alternativo del transcrito primario en la línea germinal de los machos (Chandler *et al.*, 1997).

Fuera del género *Drosophila*, el gen *tra2* se ha caracterizado en el tefrítido *Ceratitis capitata* (Salvemini *et al.*, 2009). Se transcribe en los dos sexos durante todo el desarrollo y produce un único transcrito que codifica una proteína de 251 aminoácidos con un dominio RRM y dos dominios RS. No se han encontrado evidencias de la existencia de un promotor específico de la línea germinal de los machos, por lo que parece que el mecanismo de regulación de la transcripción y el procesamiento de *tra2* en este tefrítido es

menos complejo que en *Drosophila*. Ensayos de ARN de interferencia durante el desarrollo embrionario causan una reversión de moscas genotípicamente hembras a pseudomachos fértiles, pero no tiene ningún efecto en los machos, lo que demuestra que *tra2* es esencial para el desarrollo de hembra (Salvemini *et al.*, 2009). Al igual que en *Drosophila*, la proteína Tra2, junto con la proteína Tra, presente sólo en las hembras, regulan el procesamiento específico de sexo del transcrito primario del gen *dsx*, promoviendo la inclusión de un exón específico de hembra en el ARNm y dando lugar a una proteína DsxF. Esta función la ejercería a través de su unión a las secuencias Tra-Tra2 localizadas en este exón específico de hembra. En el caso de *tra*, la unión de los complejos Tra-Tra2 a sus secuencias diana, localizadas en el pre-ARNm, previene la inclusión en el ARNm de los exones específicos de macho que contienen codones de parada de la traducción, y, por lo tanto, da lugar a una proteína Tra, funcional.

El gen *tra2* de *M. domestica* (Burghardt *et al.*, 2005) se transcribe en los dos sexos durante todo el desarrollo, y produce un único transcrito que codifica una proteína de 232 aminoácidos que contiene un dominio RRM y dos dominios RS. Al contrario de lo que ocurre en *Drosophila*, no se ha identificado ningún transcrito específico de la línea germinal de los machos. La inyección de dsARN de *tra2* en embriones de *Musca* resulta en la completa transformación de embriones genotípicamente hembras en machos fértiles. Tra2 se requiere, tanto para el procesamiento específico de hembra del pre-ARNm del gen *dsx*, como para la autorregulación del gen *F* (Burghardt *et al.*, 2005), gen clave en la determinación sexual de este insecto (Dübendorfer *et al.*, 2002) y ahora identificado como *tra* (Hediger *et al.*, 2010).

El gen *tra2* se ha identificado, también, en el lepidóptero *B. mori* (Niu *et al.*, 2005). Se expresa en las hembras y machos a lo largo de todo el desarrollo y da lugar a varios transcritos por un procesamiento alternativo que codifican seis isoformas distintas de la proteína Tra2. Todas ellas contienen un dominio RRM y dos dominios RS aunque su función se desconoce hasta el momento.

El gen *tra2* de *L. cuprina* da lugar a un único transcrito presente en ambos sexos a lo largo del desarrollo y codifica una proteína muy similar a las proteínas Tra2 de los otros dípteros (Concha y Scott, 2009).

#### 1.4.4 El gen *doublesex*

El gen *dsx* ha sido caracterizado en los dípteros *M. scalaris* (Sievert *et al.*, 1997; Kuhn *et al.*, 2000), *M. domestica* (Hediger *et al.*, 2004), *Anopheles gambiae* (el mosquito) (Scali *et al.*, 2005), *B. tryoni* (mosca de la fruta "Queensland") (Shearman y Frommer, 1998), *B. oleae* (Lagos *et al.*, 2005), *B. dorsalis* (la mosca de la fruta oriental) (Chen *et al.*, 2008), *C. capitata* (Saccone *et al.*, 2002) y en once de las especies de *Anastrepha* antes mencionadas (Ruiz *et al.*, 2007a). El gen *dsx* también ha sido aislado en el himenóptero *A. mellifera* (Cho *et al.*, 2007) y en el lepidóptero *B. mori* (Ohbayashi *et al.*, 2001; Suzuki *et al.*, 2001).

La organización molecular del gen varía entre estos insectos, pero en todos ellos *dsx* presenta un procesamiento alternativo específico de sexo, lo que da lugar a proteínas Dsx específicas de macho y hembra, que, como en el caso de *Drosophila*, comparten el extremo amino terminal y difieren en el carboxilo terminal.

#### 1.4.5 El gen *fruitless*

El gen *fru* se ha caracterizado en *D. simulans*, *D. yakuba*, *D. pseudoobscura*, *D. virilis* y *D. sukuzii* (Billeter *et al.*, 2002, *Anopheles gambiae* y *Tribolium castaneum* (Gailey *et al.*, 2006). En todos los casos, la organización molecular del gen se conserva y la proteína específica de macho FruM se genera por un procesamiento alternativo específico de sexo regulado por el complejo Tra-Tra2.

#### 1.4.6 El gen *intersex*

El gen *ix* ha sido caracterizado en los dípteros *D. virilis* y *M. scalaris* y en el lepidóptero *B. mori* (Siegal y Baker, 2005). En estos insectos, la proteína Ix parece tener una organización conservada. La región amino terminal es rica en residuos de glutamina, prolina, glicina y serina, presentes en reguladores de la transcripción. El extremo carboxilo terminal contiene una elevada proporción de aminoácidos polares y dos residuos conservados de fenilalanina que parecen ser específicos de las proteínas Ix.

El gen *ix* de *D. virilis* y *M. scalaris* recuperan la función en hembras de *D. melanogaster* mutantes para este gen, sugiriendo que la proteína Ix de estos insectos podría interaccionar con la proteína DsxF de *D. melanogaster* al igual que lo hace con su propia proteína Ix. Sin embargo, el gen *ix* de *B. mori* sólo recupera parcialmente el desarrollo sexual femenino en hembras de *D. melanogaster* mutantes para *ix*. Estos datos sugieren que las proteínas DsxF e Ix habrían coevolucionado y, que en el momento de la separación de los linajes de dípteros y lepidópteros, se habrían acumulado divergencias entre las proteínas DsxF y Ix de estas especies (Siegal y Baker, 2005).

## 1.5 MECANISMOS DE DETERMINACIÓN SEXUAL EN *Sciara*

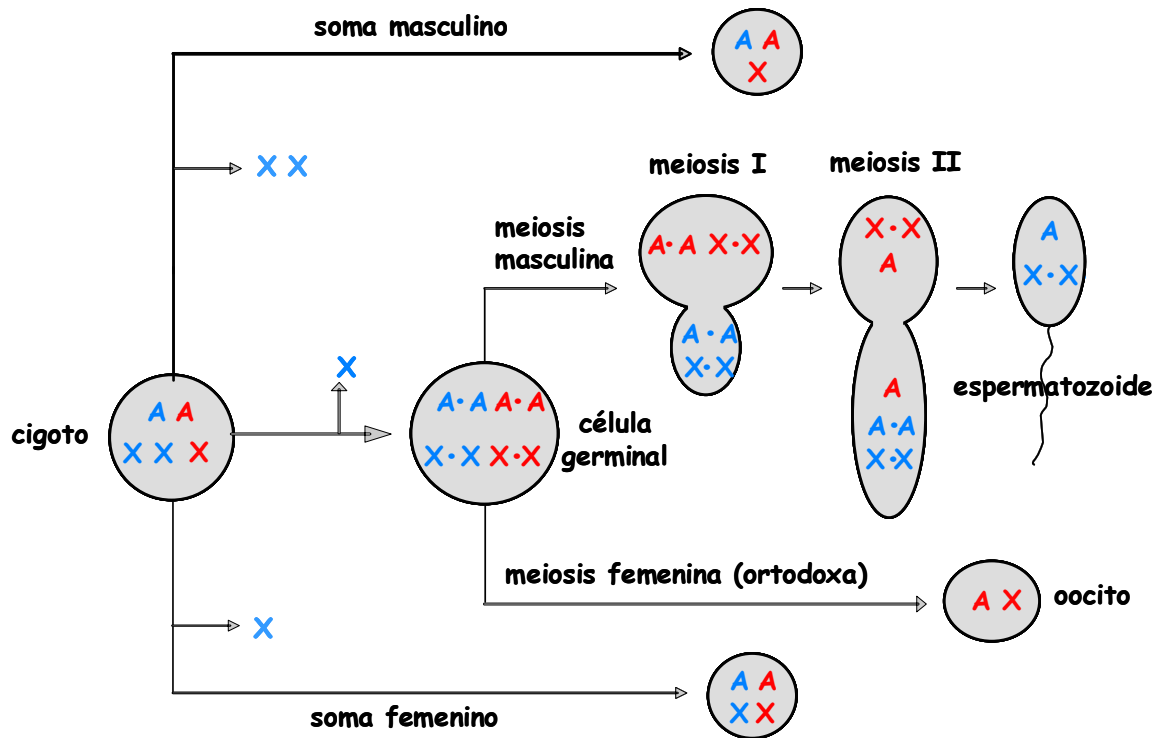
En la mayor parte de los insectos el sexo de un individuo se fija en la fecundación, tal como ocurre en *Drosophila*, *Ceratitis*, *Bactrocera* y *Musca*, siendo la constitución cromosómica de los cigotos una consecuencia directa de la constitución cromosómica de los gametos (Bull, 1983). Sin embargo, en algunos casos, las diferencias cromosómicas que determinan el sexo son la consecuencia del comportamiento especializado de los cromosomas sexuales durante los primeros estadios del desarrollo embrionario. Es el caso de los dípteros de la familia Cecidomyiidae (White, 1973; Stuart y Hatchett, 1991) y Sciaridae (DuBois, 1933; Metz, 1938), pertenecientes al suborden Nematocera.

En *Sciara* se dan los tres mecanismos básicos que producen la señal primaria que determina el desarrollo sexual que seguirá el cigoto. Estos mecanismos, que aparecen en la mayoría de las especies animales como mecanismos únicos, se dan en *Sciara* relacionados entre si y de forma concatenada: condiciones ambientales>factor materno> constitución cromosómica.

El complemento cromosómico ordinario de *Sciara* presenta tres pares de autosomas (cromosomas II, III y IV) y uno o dos cromosomas X dependiendo de si se trata de un macho o de una hembra, respectivamente. Los autosomas II y III, al igual que el cromosoma X son acrocéntricos, mientras que el cromosoma IV es metacéntrico. La especie *S. coprophila* posee, además, unos cromosomas de naturaleza heterocromática denominados cromosomas L (germline-limited), que son específicos de la línea germinal y se eliminan en las células somáticas, tanto en hembras como en machos. En *Sciara* las hembras

son 2X;2A mientras que los machos son X0;2A (Gerbi, 1986). Todos los cigotos comienzan su desarrollo con dos complementos autosómicos y tres cromosomas X (3X;2A), dos de los cuales son cromátidas hermanas de origen paterno (Metz, 1938) (los oocitos son 1X;1A y los espermatozoides 2X;1A). El ciclo cromosómico de *Sciara*, se detalla en la Figura 6. Durante las primeras divisiones embrionarias, todas las células somáticas eliminan uno o dos cromosomas X de origen paterno, lo que determinará, respectivamente, la constitución XXAA de hembras o X0AA de machos (Gerbi, 1986; Goday y Esteban, 2001). Por lo tanto, en los ciáridos, en la formación de la señal cromosómica X/A se da un proceso de impronta cromosómica que determina que los cromosomas que van a ser eliminados sean los de origen paterno. El número de cromosomas X eliminados en el embrión está controlado por un factor citoplásmico de origen materno, depositado en el oocito durante la oogénesis, y cuya naturaleza se desconoce (Metz, 1938; de Saint-Phalle y Sullivan, 1996; Sánchez y Perondini, 1999).

En etapas embrionarias más tardías, al comienzo de la segmentación de la banda germinal, todos los núcleos germinales de ambos sexos eliminan uno de los dos cromosomas X paternos. Si bien la meiosis femenina es ortodoxa, en las células germinales masculinas se produce otro evento de eliminación de cromosomas. Durante la meiosis I todo el complemento cromosómico paterno es excluido del espermatocito en una vesícula citoplásmica. En la meiosis II, las cromátidas de los autosomas segregan normalmente: la mitad de ellas son incorporadas en la vesícula citoplásmica y la otra mitad formarán el espermatozoide. El cromosoma X materno, sin embargo, no realiza la disyunción de las cromátidas hermanas, por lo que el núcleo del espermatozoide contiene dos cromosomas X de origen materno. Así, los cromosomas X paternos que son eliminados en cada generación son de origen materno, por lo que la marca de la impronta genómica, que determina los cromosomas que van a ser eliminados, es borrada en cada generación.



**Figura 6. Diagrama del ciclo cromosómico de *Sciara ocellaris*.** "A" representa los autosomas y "X" a los cromosomas X. Los cromosomas de origen materno se muestran en color rojo y los de origen paterno en color azul. En las células germinales de la meiosis masculina se representan las dos cromátidas hermanas de cada cromosoma. Nótese que en el núcleo del espermatzoide, el origen materno de los cromosomas es reconocido como paterno tras la fecundación. Modificada de Goday y Esteban (2001).

*Sciara coprophila* es una especie monogénica en la que las hembras producen en su descendencia únicamente hembras (hembras ginogénicas) o únicamente machos (hembras androgénicas) (Gerbi, 1986). La diferencia entre ambas reside en el cromosoma X: las hembras ginogénicas son X'X y las androgénicas XX (Figura 7). El cromosoma X' posee una inversión que actúa como un "aislador" que previene la formación de cromosomas recombinantes viables con su cromosoma X homólogo (Metz, 1938; Gerbi, 1986). Las hembras ginogénicas producen oocitos predeterminados a eliminar un único cromosoma X y, por lo tanto, producen solamente hembras en su descendencia. Las hembras androgénicas producen oocitos predeterminados a eliminar dos cromosomas X, produciendo solamente machos en su descendencia. De este modo, los oocitos de las hembras androgénicas y ginogénicas difieren, exclusivamente, en la cantidad de factor materno que aporta el cigoto.

*Sciara ocellaris* es una especie digénica, cuyas hembras producen en su descendencia tanto machos como hembras. Aunque a nivel de población la razón sexual (machos *versus* hembras) es 1, dicha razón varía de una hembra a otra (Mori *et al.*, 1979). Se ha observado que la razón sexual depende de la temperatura del cultivo, coincidiendo el período de sensibilidad a la temperatura con el momento de la oogénesis durante la pupa. Así, por ejemplo, a 18-20°C, el 50% de la descendencia son hembras y el otro 50% son machos; mientras que a 28-29°C, las hembras constituyen un 70% de la descendencia y los machos un 30%, no siendo dicha desviación causada por la muerte selectiva de los machos sino por la transformación de éstos en hembras; esto es, al incremento en la eliminación de uno en vez de dos cromosomas X en los embriones (Nigro *et al.*, 2007). Esto indica que a alta temperatura, las hembras producen en su gran mayoría oocitos predeterminados a eliminar uno en vez de dos cromosomas X.

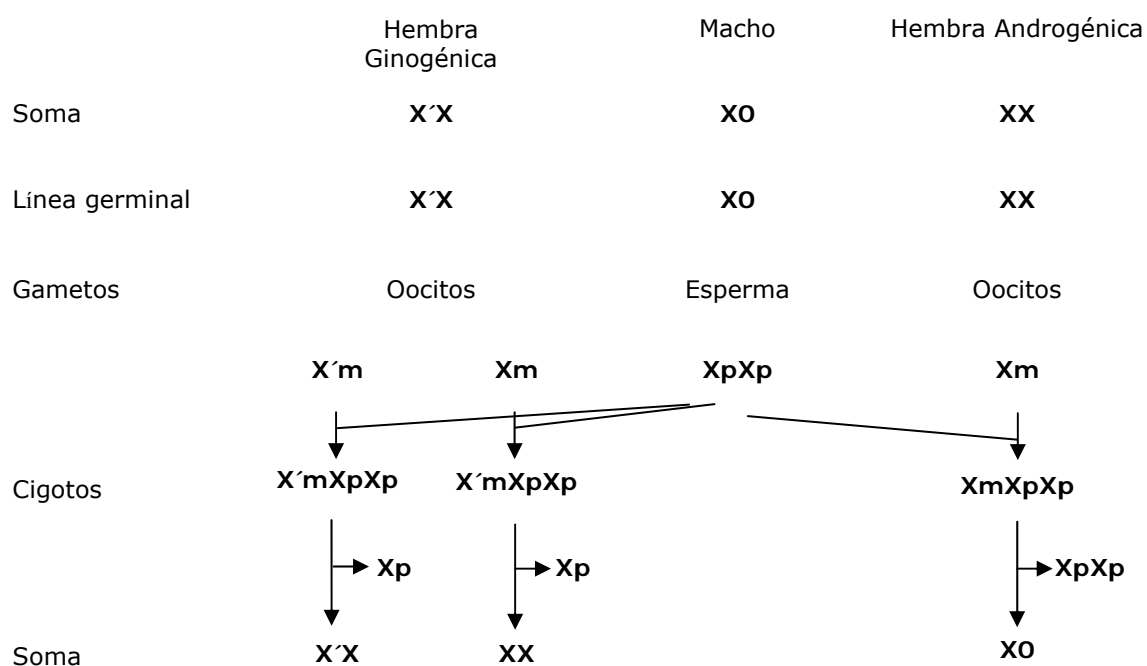


Figura 7. Producción de machos y hembras androgénicas y ginogénicas en *S. coprophila*. m representa cromosomas de origen materno. p representa cromosomas de origen paterno.

## **2. OBJETIVOS**



El trabajo se enmarca dentro del estudio de la evolución de los mecanismos de determinación sexual en insectos, utilizando *Sciara* como modelo experimental.

Uno de los propósitos es buscar en *Sciara* los genes homólogos de los que forman la cascada de la determinación sexual en *Drosophila*, y comparar cuánto se ha modificado dicha cascada entre las especies más ancestrales de los insectos y el linaje evolutivo más reciente de los drosofilidos.

En este trabajo de tesis se presenta el aislamiento y caracterización en los dípteros *S. ocellaris* y *S. coprophila* del homólogo del gen *transformer- 2* de *D.melanogaster*

### **3. MATERIALES Y MÉTODOS**

### **3.1 MATERIALES**

#### **3.1.1 Productos**

Se han utilizado productos procedentes de las marcas comerciales: Amersham Biosciences, Amersham Pharmacia Biotech, Bio Rad, Biotecx, Clontech, Invitrogen, Kodak, Merck, Perkin Elmer, Promega, Pronadisa, Qbiogen, Quiagen, Roche, GE Healthcare y Sigma.

#### **3.1.2 Medios de cultivo y soluciones**

La composición de los medios de cultivo y soluciones utilizadas en esta tesis se encuentran descritos en Maniatis *et al.*, 1982.

#### **3.1.3 Mantenimiento de *S. coprophila* y *S. ocellaris***

*S. coprophila* y *S. ocellaris* se crecieron a una temperatura de 18°C según el procedimiento indicado en Perondini y Dessen, 1985. Mediante varios cruces semanales se obtuvieron los embriones, larvas y adultos necesarios para la realización de este trabajo de tesis doctoral. Como fuente de alimento se añadió una capa de micelio triturado de champiñón comercial (Compost Villacasa, Casasimarro, Cuenca)

#### **3.1.4 Mantenimiento de *Drosophila melanogaster***

*Drosophila melanogaster* se creció a una temperatura de 25°C en botellas de cristal y tubos de plástico que contenían, aproximadamente, 20ml y 50ml, respectivamente, de medio de cultivo estándar compuesto por: 70g/l de azúcar moreno, 50g/l de harina de trigo, 100g/l de levadura de panificación, 11.5g/l de agar y 5ml/l de ácido propiónico.

## **3.2 MÉTODOS**

### **3.2.1 TÉCNICAS MOLECULARES DE ADN Y ARN**

#### **3.2.1.1 Extracción de ADN genómico**

La extracción de ADN genómico de moscas adultas de *S.coprophila*, *S.ocellaris* y *D. melanogaster* se llevó a cabo según se describe en Sambrook *et al.*, 1989. Cuando el ADN iba a ser utilizado como molde en una reacción de PCR, la extracción se realizó siguiendo el método "Quick Fly Genomic DNA prep" descrito en Berkeley Drosophila Genome Project ([www.Fruitfly.org](http://www.Fruitfly.org)).

#### **3.2.1.2 Extracción de ADN plasmídico**

La extracción de pequeñas cantidades de ADN plasmídico (5-7,5 µg) se llevó a cabo siguiendo el protocolo descrito en el kit "High Pure Plasmid Isolation Kit" de la casa comercial Roche.

Cuando fue necesario obtener cantidades mayores de ADN plasmídico (50-100 µg) se utilizó el Kit "Plasmid Midi Kit" de la casa comercial Quiagen.

#### **3.2.1.3 Electroforesis en geles de agarosa**

El ADN genómico, ADN de plásmidos y los productos de PCR se corrieron en geles de agarosa (Pronadisa) de concentraciones comprendidas entre el 0.8% y el 2%, en tampón TBE 1X (89mM Tris HCl, 89mM ácido bórico, 2mM EDTA). Como marcadores de peso molecular se utilizaron los marcadores comerciales "Molecular Weight Marker II y XIV" de Roche y "100bp DNA Step Ladder" de Promega. La electroforesis se llevó a cabo siguiendo el protocolo descrito en Maniatis *et al.*, 1982. Se incorporó a los geles Bromuro de etidio (Sharp *et al.*, 1973) y se visualizaron en un sistema "GelDoc 2000" utilizando el programa "Quantity One" de Bio-Rad.

#### **3.2.1.4 Amplificación de fragmentos de ADN mediante PCR (*Polymerase Chain Reaction*)**

Las reacciones de PCR se llevaron a cabo en un termociclador GeneAmp 2700 de Applied Biosystems. Los programas de PCR, las concentraciones de los componentes, el volumen de la reacción y las parejas de cebadores utilizadas en las diferentes amplificaciones, se especifican en el apartado correspondiente. La secuencia de los cebadores utilizados y su localización se especifican en la Tabla 1 y en la Figura 8.

#### **3.2.1.5 Clonaje de productos de PCR**

Los productos de PCR obtenidos se clonaron en el vector pCR II-TOPO según se describe en el protocolo del kit "TOPO TA Cloning" de Invitrogen o en el vector "pGEM-T Easy" de Promega siguiendo las instrucciones del fabricante. Ambos métodos permiten seleccionar las colonias con inserto a través del doble sistema Ampicilina/ $\beta$ -galactosidasa descrito en Maniatis *et al.*, 1982. La presencia del inserto en las colonias blancas se comprobó mediante PCR o por digestión con endonucleasas de restricción.

#### **3.2.1.6 Comprobación de colonias positivas por PCR**

Cada una de las colonias a analizar se resuspendió en 30 $\mu$ l de H<sub>2</sub>O. Previamente a la reacción de PCR, se calentaron 5 $\mu$ l de esta resuspensión a 100°C. A continuación, se llevó a cabo la reacción de PCR, en un volumen final de 25 $\mu$ l, utilizando las concentraciones estándar (Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,3 $\mu$ M de cada cebador y 2U de Taq polimerasa (Perkin Elmer). Se utilizaron los cebadores universales M13 Forward (-20) y M13 Reverse (-24) presentes en el vector o los cebadores específicos utilizados para amplificar el fragmento de interés.

Las condiciones de amplificación fueron las siguientes:

- Un ciclo de 4 minutos a 94°C.
- 25 ciclos, cada uno con 10 segundos a 94°C, 30 segundos a la temperatura de anillamiento de los cebadores utilizados y una fase de elongación a 72°C durante el tiempo estimado necesario para amplificar el fragmento de ADN.
- Un ciclo de 72°C durante 7 minutos.

Los productos de PCR se corrieron en geles de agarosa al 1% seleccionándose las colonias que contenían el fragmento de interés.

Los 25µl restantes de cada colonia seleccionada se crecieron toda la noche a 37°C en agitación en 7ml de LB que contenía 50µg/ml de Ampicilina.

#### **3.2.1.7 Comprobación de colonias positivas mediante “*cracking*”**

La técnica del “*cracking*” de colonias es un método rápido para comprobar la presencia de inserto en plásmidos. Se llevó a cabo siguiendo el procedimiento descrito en Maniatis *et al.*, 1982.

Los productos obtenidos se corrieron en geles de agarosa seleccionándose las colonias con inserto. Los 25µl restantes de cada colonia seleccionada se crecieron durante toda la noche a 37°C en 7 ml de medio LB que contenía 50µg/ml de Ampicilina.

#### **3.2.1.8 Comprobación de colonias positivas mediante digestión con endonucleasas de restricción**

Se digirió 1µg de ADN plasmídico con las enzimas de restricción requeridas siguiendo las instrucciones del fabricante. Los fragmentos de restricción se corrieron en geles de agarosa al 1,5%.

#### **3.2.1.9 Extracción de ARN total**

El ARN total de embriones, larvas, ovarios, testículos, abdómenes, cabeza y tórax y adultos completos de *S. ocellaris* y *S. coprophila* y de adultos de *D. melanogaster* se extrajo siguiendo el protocolo descrito en el kit "Ultraspec RNA Isolation System" de la casa comercial Biotecx. En ocasiones se utilizó para la extracción el reactivo "Trizol" (Invitrogen), siguiendo las instrucciones del fabricante.

#### **3.2.1.10 Extracción de ARN poliadenilado**

Para la obtención de ARN mensajero de adultos de *S. ocellaris*, *S. coprophila* y *D. melanogaster* se llevó a cabo según el protocolo descrito en el kit "mRNA Purification Kit" de la casa comercial GE Healthcare.

#### **3.2.1.11 Retrotranscripción (RT)**

Se llevó a cabo a partir de ARN total o ARN mensajero, bien utilizando cebadores específicos o con un *oligo-dT* tal y como se describe en el protocolo de la enzima "SuperScript<sup>™</sup> II RNase H Reverse Transcriptase" (Invitrogen). El cebador utilizado, el tipo de ARN y las cantidades correspondientes se especifican en cada apartado. Antes de realizar la retrotranscripción siempre se comprobó que no existía contaminación del ARN de partida con ADN genómico, realizando una PCR sobre el mismo.

#### **3.2.1.12 Secuenciación de ADN**

Todas las reacciones de secuenciación de esta Tesis han sido realizadas en el Servicio de Secuenciación Automática de ADN del Centro de Investigaciones Biológicas ([www.secugen.es](http://www.secugen.es)).

### **3.2.1.13 Análisis de las secuencias**

El análisis de todas las secuencias de ADN y proteínas se han realizado utilizando los siguientes programas informáticos: DNA compare (Enterlist, Compare, Translate y Restrict) Editview y BioEdit y las siguientes bases de datos: NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)), BLAST ([www.ncbi.nlm.nih.gov/BLAST](http://www.ncbi.nlm.nih.gov/BLAST)), FlyBase (<http://flybase.org/>) CLUSTALW ([www.ebi.ac.uk/clustalw](http://www.ebi.ac.uk/clustalw)) y FASTA ([www.ebi.ac.uk/Tools/fasta33/index.htm](http://www.ebi.ac.uk/Tools/fasta33/index.htm)).

### **3.2.1.14 Análisis del posible sitio de inicio de la transcripción**

Para analizar los posibles sitios putativos de inicio de la transcripción de los transcritos del gen *tra2* de *S. ocellaris* y *S. coprophila* se utilizó el programa informático: "Neural Network Promoter Prediction" ([www.fruitfly.org/seq\\_tools/promoter.html](http://www.fruitfly.org/seq_tools/promoter.html)).

### **3.2.1.15 Diseño de los oligonucleótidos.**

El diseño de los oligonucleótidos sintéticos utilizados como cebadores en esta tesis se realizó utilizando los programas: "Primer 3" ([http://biotools.umassmed.edu/bioapps/primer3\\_www.cgi](http://biotools.umassmed.edu/bioapps/primer3_www.cgi)) y "Oligo Analyzer 3.1" (<http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/>). La síntesis de todos los oligonucleótidos se llevó a cabo por la casa comercial Roche.

## **3.2.2 AISLAMIENTO Y CARACTERIZACIÓN DEL GEN *tra2* DE *S. ocellaris***

### **3.2.2.1 RT-PCR con oligonucleótidos degenerados**

La reacción de retrotranscripción se realizó a partir de 5µg de ARN total de hembras adultas de *S. ocellaris* utilizando como cebador un *oligo-dT*. Para la amplificación por PCR se empleó 1µl de ADNc, en un volumen final de 25µl, utilizando las siguientes condiciones: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Perkin Elmer). Como cebadores se utilizaron dos parejas de oligonucleótidos sintéticos degenerados "MAR26" y "MAR17", para la primera amplificación y "MAR26" y "TRA2B" para la



reamplificación, todos ellos correspondientes a las regiones conservadas del dominio RRM de los genes *tra2* de *M. domestica*, *D. virilis* y *D. melanogaster*.

El programa de amplificación fue el siguiente:

- Un primer ciclo de 2 minutos a 94°C.
- 45 ciclos, cada uno de ellos con 30 segundos a 94°C, 1 minuto a 42°C y 10 segundos a 72°C.
- una elongación final de 7 minutos a 72°C.

1 µl del producto del PCR anterior se reamplificó, utilizando las mismas condiciones, con el siguiente programa de amplificación:

- Un primer ciclo de 2 minutos a 94°C.
- 45 ciclos, cada uno de ellos con 30 segundos a 94°C, 1 minuto a 65°C y 10 segundos a 72°C.
- una elongación final de 7 minutos a 72°C

Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

### 3.2.2.2 RACE ("*Rapid Amplification of cDNA Ends*")

La caracterización de los extremos 5' y 3' del ADNc del gen *tra2* de *S. ocellaris* se llevó a cabo mediante la técnica de RACE. Para la construcción de la genoteca, se utilizó 1µg de ARN poliadenilado de hembras adultas de *S. ocellaris*. La síntesis de la primera y segunda cadena de ADNc, la ligación de los adaptadores y las reacciones de PCR 3' y 5' RACE se llevaron a cabo tal y como se describe en el kit "Marathon<sup>TM</sup> cDNA Amplification" (Clontech)

Se realizaron dos rondas de amplificación utilizando como cebadores, en ambos casos, el oligonucleótido del adaptador AP1 (en la primera PCR) o AP2 (en la segunda PCR), suministrados en el kit, y un oligonucleótido específico diseñado sobre la secuencia obtenida del gen *tra2* de *S. ocellaris*: "*RACE5' SoT2*" (para caracterizar el extremo 5') y "*RACE3' SoT2.2*" (para caracterizar el extremo 3'). Para las reacciones de PCR se utilizó la enzima "Advantage 2 PCR Enzyme System" (Clontech).

Los productos de amplificación fueron analizados por electroforesis en geles de agarosa al 0,8%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

### 3.2.2.3 *Genome Walker* (Paseo Cromosómico)

Para llevar a cabo el paseo cromosómico se empleó el kit "BD Genome Walker™ Universal kit" (Clontech). La síntesis de la librería genómica, la ligación de los adaptadores y la PCR se llevó a cabo siguiendo el protocolo descrito en el kit. Para las reacciones de PCR se utilizó la enzima "Advantage 2 PCR Enzyme System" (Clontech).

En cada reacción se utilizó una pareja de cebadores formada por el oligonucleótido "AP1" suministrado en el kit y un oligonucleótido específico correspondiente a la secuencia del gen *tra2* de *S. ocellaris*.

Los cebadores específicos que se utilizaron fueron los siguientes: "GWtra2-1s", "GWtra2-3s" y "GW5´SoT2".

Los productos de la amplificación fueron analizados por electroforesis en geles de agarosa al 0,8%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

### 3.2.2.4 PCR de largo alcance ("*long PCR*")

La técnica de "*long PCR*" permite amplificar un producto de miles de pares de bases, más extenso que mediante una PCR convencional.

Se emplearon 100ng de ADN genómico de *S. Ocellaris*, en un volumen final de reacción de 25µl, y las siguientes condiciones: Buffer 1x, 1.75mM de MgCl<sub>2</sub>, 0,35mM de dNTPs, 0.3µM de cada cebador y 5U de la enzima "Expand Long Template Enzyme Mix" (Roche). Los cebadores utilizados fueron: "5´SoT2" y "I5SoT2".

El programa de amplificación fue el siguiente:

- Un ciclo inicial de 5 minutos a 95°C.
- 45 ciclos, cada uno de ellos con 30 segundos a 94°C, 30 segundos a 58°C y 4 minutos a 68°C.

- Una elongación final de 7 minutos a 68°C.

Posteriormente, 0,5µl, de una dilución 1:100 del producto anterior, se reamplificaron, utilizando los mismos cebadores y las mismas condiciones.

Los productos de amplificación fueron analizados por electroforesis en geles de azarosa al 0,8%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

### **3.2.2.5 Expresión del gen *tra2* de *S.ocellaris* durante el desarrollo mediante RT-PCR**

La retrotranscripción se llevó a cabo a partir de 5µg de ARN total extraído de una mezcla de embriones machos y hembras de 24 horas, una mezcla de larvas machos y hembras de 28 días, cabeza y tórax de machos, cabeza y tórax de hembras, abdómenes de machos, abdómenes de hembras y adultos de ambos sexos, testículos y ovarios de *S. ocellaris* utilizando el cebador "T30". Los cebadores que se emplearon en la reacción de PCR fueron: "*expSoT2.1*", localizado en el exón 5 y un cebador perteneciente al exón 6: "*expSoT2.2*". Como control interno se amplificó un fragmento de la proteína ribosómica Rpl10 utilizando los cebadores "*rpl10.1*" y "*rpl10.2*".

Para la amplificación por PCR se utilizó, en los tres casos, 1µl de ADNc, en un volumen final de 25µl, con las siguientes condiciones: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Perkin Elmer).

Para la amplificación del transcrito y la proteína ribosómica *Rpl10* el programa utilizado fue el siguiente:

- Un ciclo inicial de 4 minutos a 94°C.
- 55 ciclos cada uno de ellos con 15 segundos a 94°C, 15 segundos a 58°C y 15 segundos a 72°C.
- Una elongación final de 7 minutos a 72°C.

Los productos de la amplificación fueron analizados por electroforesis en geles de agarosa al 0,8%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

### 3.2.2.6 Expresión del transcrito *Sotra2-204*

La presencia del transcrito *Sotra2-204* se analizó mediante RT-PCR. La retrotranscripción se llevó a cabo a partir de 5µg de ARN total de hembras y de machos por separado de *S. ocellaris*. Para la amplificación por PCR se utilizaron los cebadores "M1So.1" y "expSoT2.2".

Se utilizó 1µl de ADNc, en un volumen final de 25µl, con las siguientes condiciones: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Perkin Elmer).

El programa de amplificación fue el siguiente:

- Un ciclo inicial de 4 minutos a 94°C.
- 40 ciclos, cada uno de ellos con 15 segundos a 94°C, 20 segundos a 58°C y 30 segundos a 72°C.
- Un ciclo de elongación de 7 minutos a 72°C.

Los productos de amplificación fueron analizados por electroforesis en geles de agarosa al 0,8%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

### 3.2.2.7 Expresión del transcrito *Sotra2-130*

La presencia del transcrito *Sotra2-130* se analizó mediante RT-PCR. La retrotranscripción se llevó a cabo a partir ARN total de cabeza y tórax de machos, cabeza y tórax de hembra, testículos y ovarios de *S. ocellaris*. Para la amplificación por PCR se utilizaron los cebadores "*S. ocel*3.1" y "expSoT2.2".

Se utilizó 1µl de ADNc, en un volumen final de 25µl, con las siguientes condiciones: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Perkin Elmer).

El programa de amplificación fue el siguiente:

- Un ciclo inicial de 4 minutos a 94°C.
- 35 ciclos, cada uno de ellos con 15 segundos a 94°C, 20 segundos a 56°C y 30 segundos a 72°C.
- Un ciclo de elongación de 7 minutos a 72°C.

Los productos de amplificación fueron analizados por electroforesis en geles de agarosa al 0,8%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

#### **3.2.2.8 Cuantificación de la expresión de los transcritos *SoTra2-251* y *SoTra2-130* en *S. ocellaris***

La expresión diferencial de los transcritos *SoTra2-251* y *SoTra2-130* se analizó mediante RT-PCR. La retrotranscripción se llevó a cabo a partir de ARN total de cabeza y tórax de machos, cabeza y tórax de hembra, testículos y ovarios de *S. ocellaris* utilizando como cebador un “oligo dT”.

Para la amplificación por PCR se utilizó, en los tres casos, 1µl de ADNc, en un volumen final de 25µl, con las siguientes condiciones: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Perkin Elmer). Los cebadores empleados fueron “*M1So.1*” y “*expSoT2.2*”.

Para definir las condiciones óptimas del experimento se realizaron tres programas de amplificación diferentes que únicamente variaban en el número de ciclos (30, 35 ó 40).

El programa elegido fue el siguiente:

- Un ciclo inicial de 4 minutos a 94°C.
- 35 ciclos, cada uno de ellos con 15 segundos a 94°C, 20 segundos a 62°C y 30 segundos a 72°C.
- Un ciclo de elongación de 7 minutos a 72°C.

Se realizaron tres réplicas.

Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1.5%. La concentración correspondiente a cada banda amplificada se midió en un equipo GelDoc 2000 utilizando el programa “QuantityOne” (Bio-Rad).

### **3.2.3 AISLAMIENTO Y CARACTERIZACIÓN DEL GEN *tra2* DE *S. coprophila***

#### **3.2.3.1 Amplificación mediante PCR de un fragmento de ADN genómico**

Se utilizaron 100ng de ADN genómico de *S. coprophila* en un volumen final de PCR de 25µl. Las condiciones de PCR utilizadas fueron las estándar: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0.2µM de cada uno de los cebadores y 2U Taq polimerasa (Invitrogen). Los cebadores que se utilizaron fueron "*expSoT2.1*" y "*expSoT2.2*" localizados en los exones E5 y E6, respectivamente, del gen *tra2* de *S. ocellaris*.

El programa de amplificación fue el siguiente:

- Un ciclo inicial de 4 minutos a 94°C.
- 45 ciclos, cada uno de ellos con 30 segundos a 94°C, 30 segundos a 54°C y 1 minuto a 72°C.
- Un ciclo de elongación de 7 minutos a 72°C.

Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1%, clonados en el vector pGEM-T easy (Promega) y secuenciados con los oligonucleótidos universales M13 Forward (-47) y M13 Reverse (-48).

#### **3.2.3.2 RACE ("*Rapid Amplification of cDNA Ends*")**

La caracterización de los extremos 5' y 3' de los ADNc de macho y hembra del gen *tra2* de *S. coprophila* se llevó a cabo mediante RACE. Para la construcción de las genotecas se partió de 1µg de ARN poliadenilado de machos y hembras adultos de *S. coprophila*, separadamente. La síntesis de la primera y segunda cadena de ADNc, la ligación de los adaptadores y las reacciones de PCR 3' y 5' RACE se llevaron a cabo tal como se describe en el kit "Marathon<sup>TM</sup> cDNA Amplification" (Clontech).

Se realizaron dos rondas de amplificación, utilizando como cebadores, en ambos casos, el oligonucleótido del adaptador AP1 (en la primera PCR) o AP2 (en la segunda PCR), suministrados en el kit y un oligonucleótido específico diseñado sobre la secuencia previamente obtenida del gen *tra2* de *S. coprophila*. Los cebadores específicos que se utilizaron son: "RACESco5´2" (para caracterizar el extremo 5') "RACE Sco 3´" (para caracterizar el extremo 3'). Para las reacciones de PCR se utilizó la enzima "Advantage 2 PCR enzyme system" (Clontech).

Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1%, clonados en el vector pGEM-T easy (Promega) y secuenciados con los oligonucleótidos universales M13 Forward (-47) y M13 Reverse (-48).

### 3.2.3.3 *Genome Walker* (Paseo Cromosómico)

Para llevar a cabo el paseo cromosómico se empleó el kit "BD Genome Walker™ Universal kit" (Clontech). El procedimiento utilizado fue el mismo que para el caso de *S. ocellaris* (ver apartado 2.2.3).

Los oligonucleótidos que se utilizaron en este caso fueron: "RACESco5´", "RACE Sco 3´", "GWScop5´", "GWScop5´2", "ScopE2+GW" y "ScopE2-GW".

### 3.2.3.4 RT-PCR con cebadores específicos del gen *Sctra2*

La retrotranscripción se realizó a partir de 5µg de ARN total de testículos y hembras adultas de *S. coprophila* utilizando el cebador "T30". Para la amplificación por PCR se utilizó 1µl de ADNc, en un volumen final de 25µl, usando las condiciones estándar: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0.2µM de cada uno de los cebadores y 2U Taq polimerasa (Invitrogen). Los cebadores utilizados fueron "Scop5´" y "Scop3´", localizados en los exones E1 y E6, respectivamente.

El programa de amplificación utilizado fue:

- Un ciclo inicial de 4 minutos a 94°C.
- 55 ciclos, cada uno de ellos con 20 segundos a 94°C, 20 segundos a 58°C y 1 minuto a 72°C.
- Un ciclo de elongación de 7 minutos a 72°C.

Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1.2%, clonados en el vector pGEM-T easy (Promega) y secuenciados con los oligonucleótidos universales M13 Forward (-47) y M13 Reverse (-48).

### **3.2.3.5 Expresión del gen *tra2* de *S. coprophila* durante el desarrollo mediante RT-PCR**

La retrotranscripción se llevó a cabo a partir de 5µg de ARN total de embriones machos y hembras de 24 horas, larvas machos y hembras de 28 días, adultos de ambos sexos, cabezas y tórax de ambos sexos, testículos y ovarios de *S. ocellaris* utilizando como cebador el oligo "T30". Los cebadores que se utilizaron en la reacción de PCR fueron "expSoT2.1" y "expSoT2.2", localizados en los exones E5 y E6, respectivamente.

Como control interno se amplificó un fragmento de la proteína ribosómica *Rpl10* utilizando los cebadores "rpl10.1" y "rpl10.2".

Para la amplificación por PCR se utilizó, en ambos casos, 1µl de ADNc, en un volumen final de 25µl, con las siguientes condiciones: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Perkin Elmer).

Se realizaron dos programas de amplificación diferentes.

Para la expresión del gen *Sctra2*:

- Un ciclo inicial de 4 minutos a 94°C.
- 55 ciclos cada uno de ellos con 20 segundos a 94°C, 15 segundos a 56°C y 15 segundos a 72°C.
- Una elongación final de 7 minutos a 72°C.

Para la expresión del gen control *rpl10*:

- Un ciclo inicial de 4 minutos a 94°C.
- 55 ciclos cada uno de ellos con 20 segundos a 94°C, 20 segundos a 58°C y 20 segundos a 72°C.
- Una elongación final de 7 minutos a 72°C.



Los productos de la amplificación fueron analizados por electroforesis en geles de agarosa al 0,8%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

### **3.2.4 MOSCAS TRANSGÉNICAS DE *Drosophila melanogaster***

Para generar moscas transgénicas de *D. melanogaster* el correspondiente gen reportero se clonó en el vector pUAST (Brand y Perrimon, 1993). Este vector se caracteriza por contener cinco secuencias UAS reconocidas por el factor de transcripción GAL-4, de modo que el transgén solo se expresará si dicho factor está presente. El vector pUAST contiene además un marcador  $w^+$  que confiere color de ojos rojos a las moscas transgénicas.

La información de las construcciones GAL-4, así como de los alelos mutantes se encuentran descritas en Lindsley y Zimm, 1992 y en FlyBase ([www.flybase.org](http://www.flybase.org)).

*TDSO* hace referencia a UAS::Sotra2-ADNc.

#### **3.2.4.1 Construcción del elemento transgénico *TDSO***

Para realizar la construcción *TDSO*, utilizada en la transformación de *D. melanogaster* para obtener las líneas transgénicas, se introdujo un fragmento de ADNc de 920pb de *S. ocellaris* en el sitio EcoRI del vector pUAST. Este fragmento contenía el marco abierto de lectura completo del gen *tra2* y fue amplificado por RT-PCR. Para la retrotranscripción se utilizaron 5µg de ARN total de hembras adultas de *S. ocellaris* y como cebador el *oligo-dT*. Los cebadores utilizados para la amplificación por PCR fueron los oligonucleótidos "5´SoT2" y "3´SoT2".

En la reacción de PCR se utilizó 1µl del ADNc, en un volumen final de 25µl, con las siguientes condiciones: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Perkin Elmer).

El programa de amplificación utilizado fue el siguiente:

- Un ciclo inicial de 5 minutos a 95°C.
- 45 ciclos, cada uno de ellos con 30 segundos a 94°C, 30 segundos a 58°C y 1 minuto a 72°C.
- Un ciclo de 7 minutos a 72°C.

Los productos de la amplificación fueron analizados por electroforesis en geles de agarosa al 1%, clonados en el vector pCR 2.1- Topo (Invitrogen) y secuenciados con los oligonucleótidos universales M13 (-20) y M13 reverse (-24).

El fragmento de ADNc de *S. ocellaris* se liberó del vector "pCR 2.1 – Topo" (Invitrogen), en el que se había clonado previamente, mediante digestión con la enzima de restricción EcoRI. El producto de la digestión se corrió en un gel de agarosa al 0, 8% y el fragmento de interés fue eluido del gel siguiendo el protocolo descrito en el kit "GeneClean Turbo" (Qbio gene).

El inserto y el vector, previamente digerido con la enzima de restricción EcoRI, se ligaron en las cantidades y proporciones especificadas en el kit "DNA Rapid Ligation" (Roche) y se transformaron bacterias *E. coli* DH5 $\alpha$ . Las colonias con inserto se seleccionaron mediante PCR utilizando las mismas condiciones y cebadores descritos anteriormente.

Se extrajo el ADN de las colonias seleccionadas y se secuenció para comprobar si el inserto se había clonado en la orientación deseada.

#### **3.2.4.2 Generación de moscas transgénicas**

La producción de las moscas de *D. melanogaster* transgénicas fue realizada por *Genetic Services* (Sudbury, MA, USA).

#### **3.2.4.3 Localización cromosómica de los transgenes**

La localización cromosómica de los transgenes, en las diferentes líneas de moscas transgénicas, se comprobó mediante los siguientes cruces:

**Cruce 1.** Machos  $w^+$  (ojo rojo- naranja) de cada una de las líneas transgénicas se cruzaron con hembras  $yw$ ;  $CyO$ ,  $Cy/ T$  (2; 3)  $Xa/ MKRS$ ,  $Sb$ . Si en la descendencia

de este cruce (F1) todos los machos son de fenotipo  $w$  (ojos blancos), se deduce que el transgén se ha insertado en el cromosoma X. Sin embargo, si en la descendencia aparecen también machos de fenotipo  $w^+$ , se deduce que el transgén se ha insertado en un autosoma. En este último caso, se realizó un segundo cruce para comprobar si la inserción se encontraba en el cromosoma II o en el cromosoma III.

**Cruce 2.** Machos de fenotipo  $w^+ CyO Sb$ , procedentes del cruce 1 se cruzaron con hembras  $yw$ . Si en la descendencia de este cruce aparecen moscas de fenotipo  $w^+ Cy$ , se deduce que el transgén se ha insertado en el cromosoma III, mientras que, si aparecen en la descendencia moscas de fenotipo  $w^+ Sb$ , se deduce que el transgén se ha insertado en el cromosoma II. En ambos casos, la descendencia de este cruce se autocruzó para balancear el transgén, bien sobre  $CyO$ ,  $Cy$ , si la inserción se había producido en el cromosoma II, o bien sobre  $MKRS$ ,  $Sb$ , si la inserción se había producido en el cromosoma III.

#### 3.2.4.4 Comprobación de la presencia del transgén *TDSO* en moscas de *D. melanogaster* transgénicas

La presencia del transgén *TDSO* se comprobó mediante PCR, a partir de 100 ng de ADN genómico de adultos de ambos sexos (ver apartado 2.1.1) utilizando un cebador del vector "5' *pUAST*" y un cebador del inserto "3' *SoT2*".

Para la amplificación por PCR se empleó un volumen final de 25µl, usando las siguientes condiciones: Buffer 1X, 1.5mM de  $MgCl_2$ , 0,2mM de dNTPs, 0,2µM de cada cebador y 2U de Taq polimerasa (Invitrogen).

El programa de amplificación utilizado fue el siguiente:

- Un ciclo inicial de 5 minutos a 94°C.
- 50 ciclos cada uno de ellos con 30 segundos a 94°C, 30 segundos a 58°C y 1 minuto a 72°C.
- Un ciclo de 7 minutos a 72°C.

Los productos de amplificación fueron analizados en geles de agarosa al 1%, clonados en el vector pGEM-T easy (Promega) y secuenciados con los oligonucleótidos universales M13 Forward (-47) y M13 Reverse (-48).

### 3.2.4.5 Análisis de la expresión de los transgenes

El análisis de la expresión del transgén *TDSO* en las líneas #1, #2, #3 y #8 se llevó a cabo mediante RT-PCR sobre moscas transgénicas de *D. melanogaster* a las que se les dio tres choques térmicos (37°C), de tres horas cada uno, a lo largo de 36 horas, manteniéndose las moscas a 25°C entre los choques térmicos. Las moscas control no recibieron choque térmico.

La retrotranscripción se llevó a cabo a partir de 5µg de ARN total utilizando el cebador "T30". Para la amplificación por PCR se empleó, en ambos casos, 1µl de ADNc en un volumen final de 50µl. Los cebadores utilizados fueron "5´*SoT2.1*" y "3´*SoT2.2*", que amplifican un fragmento de 920pb. Como control se analizó la expresión del gen *rp49* que codifica la proteína ribosómica 49 de *D. melanogaster* (Ramos-Onsins et al., 1998). Se utilizaron los cebadores "*rp49.1*" y "*rp49.2*", que amplifican un fragmento de 286pb.

Las condiciones fueron las siguientes: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0.2µM de cada uno de los cebadores y 2U Taq polimerasa (Invitrogen).

El programa de amplificación fue el siguiente:

- Un ciclo de 4 minutos a 94°C.
- 55 ciclos, cada uno con 20 segundos a 94°C, 20 segundos a 58°C y 25 segundos a 72°C.
- Un ciclo final de 7 minutos a 72°C.

Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1.2%.

### 3.2.4.6 Efecto del transgén *TDSO* sobre el procesamiento del pre-ARNm del gen *doublesex* de *Drosophila melanogaster*.

Se seleccionaron hembras vírgenes de cuatro líneas transgénicas *TDSO* y se cruzaron con machos de una línea de *D. melanogaster* que contiene el factor de transcripción GAL-4 (*w; Hs-GAL4 / CyO,Cy*) que sólo expresa el factor de transcripción tras un choque térmico. Los cruces se mantuvieron a 25°C.

A las moscas descendientes de cada uno de estos cruces se les dividió en dos grupos. Uno de ellos se mantuvo a 25°C. Este grupo son las moscas control, ya que no expresan el factor de transcripción GAL-4, y, por tanto, tampoco el

transgén. Al otro grupo se le dio tres choques térmicos (37°C) de tres horas cada uno a lo largo de 36 horas, manteniéndose las moscas a 25°C entre los choques térmicos. Estas moscas van a expresar el factor de transcripción GAL-4 y por lo tanto el transgén.

La expresión de *dsx* se analizó mediante RT-PCR. La retrotranscripción se llevó a cabo a partir de 5µg de ARN total de moscas adultas de genotipo *yw/w; df(2R)trix/tra2<sup>B</sup>; TDSO#/HS-Gal4*, con utilizando el cebador "T30".

Para la detección del transcrito *dsxF* endógeno de *D. melanogaster*, se utilizaron como cebadores en la reacción de PCR los oligonucleótidos "*dsx1*" y "*dsxF1*", amplifican un fragmento de 654pb del transcrito específico de hembra de *dsx*. Como control interno se analizó la expresión del gen *rp49*, descrita en el apartado anterior.

Para la amplificación por PCR se utilizó, en ambos casos, 1µl de ADNc en un volumen final de 50µl. Las condiciones fueron las siguientes: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0.2µM de cada uno de los cebadores y 2U Taq polimerasa (Invitrogen).

El programa de amplificación utilizado fue el siguiente:

- Un ciclo inicial de 5 minutos a 94°C.
- 40 ciclos, cada uno de ellos con 15 segundos a 94°C, 20 segundos a 58°C y 20 segundos a 72°C.
- Un ciclo final de 7 minutos a 72°C.

Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1.2%, clonados en el vector pGEM-T easy (Promega) y secuenciados con los oligonucleótidos universales M13 Forward (-47) y M13 Reverse (-48).

Para la detección del transcrito *dsxM* endógeno de *D. melanogaster*, se utilizaron como cebadores en la reacción de PCR los oligonucleótidos "*dsx1*" y "*dsx.M3*", que amplifican un fragmento de 481pb del transcrito específico de macho de *dsx*. Como control interno se analizó la expresión del gen *rp49*, descrita anteriormente.

Para la amplificación por PCR se utilizó, en ambos casos, 1µl de ADNc en un volumen final de 50µl. Las condiciones fueron las siguientes: Buffer 1x, 1.5mM de MgCl<sub>2</sub>, 0.2 mM de dNTPs, 0.2µM de cada uno de los cebadores y 2U Taq polimerasa (Invitrogen).

El programa de amplificación utilizado fue el siguiente:

- Un ciclo inicial de 5 minutos a 94°C.
- 50 ciclos, cada uno de ellos con 15 segundos a 94°C, 20 segundos a 60°C y 20 segundos a 72°C.
- Un ciclo final de 7 minutos a 72°C.

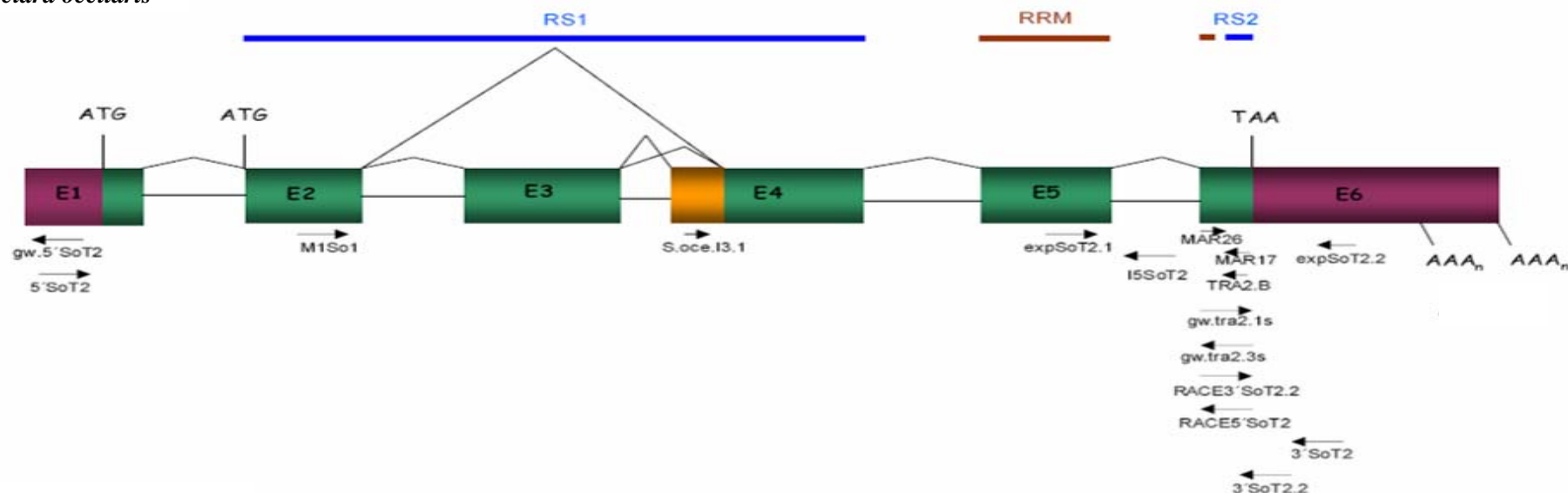
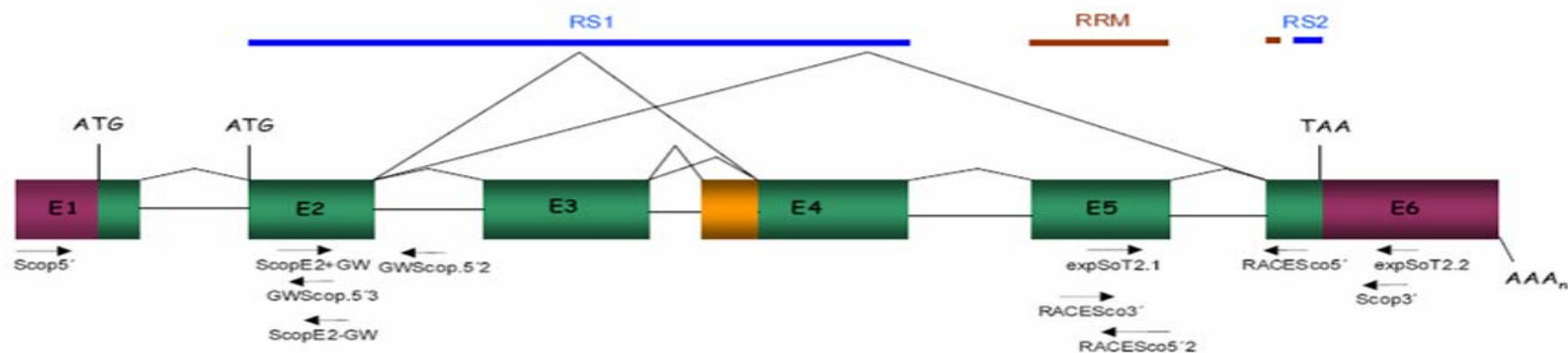
Los productos de PCR fueron analizados por electroforesis en geles de agarosa al 1.2%, clonados en el vector pGEM-T easy (Promega) y secuenciados con los oligonucleótidos universales M13 Forward (-47) y M13 Reverse (-48).

CEBADOR ▼	SECUENCIA ▼
<b>MAR 26</b>	➤ 5' MGI AGY CGN GGN TTY TGY TTY 3'
<b>MAR 17</b>	➤ 5' GTR TGI GSI CGY TKN GTD ATN GA 3'
<b>TRA2B</b>	➤ 5' NCK RTC RTC DAT YTC CAT NCC 3'
<b>GWtra2-1S</b>	➤ 5' TCG TCT ACT TTG ATC GCA CCG ATG AC 3'
<b>GWtra2-3S</b>	➤ 5' TTA TCG ATT TCC ATG CCG GTG CAA TGG 3'
<b>RACE5´SoT2</b>	➤ 5' ACG ACC GTC ATT TCG ATC TCT TTC GTG G 3'
<b>RACE3´SoT2.2</b>	➤ 5' CGC CGT CGT TAC AGT AGA TCA CGT TCG 3'
<b>GW5´SOT2</b>	➤ 5' CTG GAG ATG TAT TTA TAG CGG TTC TTG 3'
<b>I5SoT2</b>	➤ 5' GGC TAC AGG AAG TCC CGT TC 3'
<b>5´SoT2</b>	➤ 5' ATA AAA GCG AGC CAA GAA CC 3'
<b>3´SoT2</b>	➤ 5' ATT TTG TGG TCC CGC TAT CT 3'
<b>3´SoT2.2</b>	➤ 5' CCC GCG AAA AAT TAC TTC 3'
<b>expSoT2.1</b>	➤ 5' GTT GCC TTG GTG TGT TTG GT 3'
<b>expSoT2.2</b>	➤ 5' CTT TGT AAA TGC CCG GTG TT 3'
<b>M1So.1</b>	➤ 5' CCG TCG GCA GGA AGT ACA 3'
<b>S.ocol3.1</b>	➤ 5' CTC GAA GCA ACC ACA AGC 3'
<b>dsx1</b>	➤ 5' CAT CGG GAA CAT CGG TGA TC 3'
<b>dsxF1</b>	➤ 5' AAC GGC TGT CAC ACA TCG AC 3'
<b>dsxM3</b>	➤ 5' GGT GGA CAA ATC TGT GTG AG 3'
<b>RACE5co5´</b>	➤ 5' GCC GCA ACG TTA TTG AAT AGT CAA CCC G 3'
<b>RACE3co3´</b>	➤ 5' GAG TGT CTA CAC CAC GAA AGA TCA AAT C 3'
<b>RACE5co5´2</b>	➤ 5' GCA TCG ACC ACA ACC TGA ACG TGC TCG A 3'
<b>GWScop5´3</b>	➤ 5' GTA AAC CTG ATA TCT GTT CCG AGT TCT 3'
<b>GWScop5´2</b>	➤ 5' CTG ACT TGG TCA GGT GAT AGA AAA AC 3'

<b>ScopE2+GW</b>	➤	5' GTT CGA GAA GTC GGA GTT ATG AAG TT 3'
<b>ScopE2-GW</b>	➤	5' TGT GAA CTT CAT AAC TCC GAC TTC TC 3'
<b>Scop5'</b>	➤	5' AAG CCA CGC TTG AAA GTG AT 3'
<b>Scop3'</b>	➤	5' CAT CGG TAA GGC AAA CGA AC 3'
<b>rpl10.1</b>	➤	5' GAG TGC ATC CAT TCC ACG T 3'
<b>rpl10.2</b>	➤	5' GTA TTT GAC GTT GCA ACC GT 3'
<b>rp49.1</b>	➤	5' ATC CGC CAC CAG TCG GAT C 3'
<b>rp49.2</b>	➤	5' TGG CGC GCT CGA CAA TCT C 3'
<b>Oligo d(T)</b>	➤	5' TGC CAC GCT CGA CTA GTA CGT <sub>N22</sub> 3'
<b>T30</b>	➤	5' T <sub>N30</sub> 3'
<b>AP1</b>	➤	5' CCA TCC TAA TAC GAC TCA CTA TAG GGC 3'
<b>AP2</b>	➤	5' ACT CAC TAT AGG GCT CGA GCG GC 3'
<b>M13 Forward</b>	➤	5' GTA AAA CGA CGG CCA G 3'
<b>M13 Reverse</b>	➤	5' CAG GAA ACA GCT ATG A 3'
<b>pUAST 5'</b>	➤	5' CAA ATA AAC AAG CGC AGC TG 3'

**Tabla 1.** Secuencia de los cebadores empleados en esta tesis.



*Sciara ocellaris**Sciara coprophila*

**Figura 8.** Localización en el gen *tra2* de los cebadores empleados en *S. ocellaris* y en *S. coprophila*. Los rectángulos coloreados representan los exones. Las regiones 5'UTR y 3'UTR se representan en color morado. Las líneas negras representan los distintos tipos de procesamiento alternativo así como los intrones. Se indican los codones de inicio y de parada de la traducción y los sitios de poliadenilación. Las líneas gruesas indican los dominios RS (azul) y el dominio RRM (marrón).

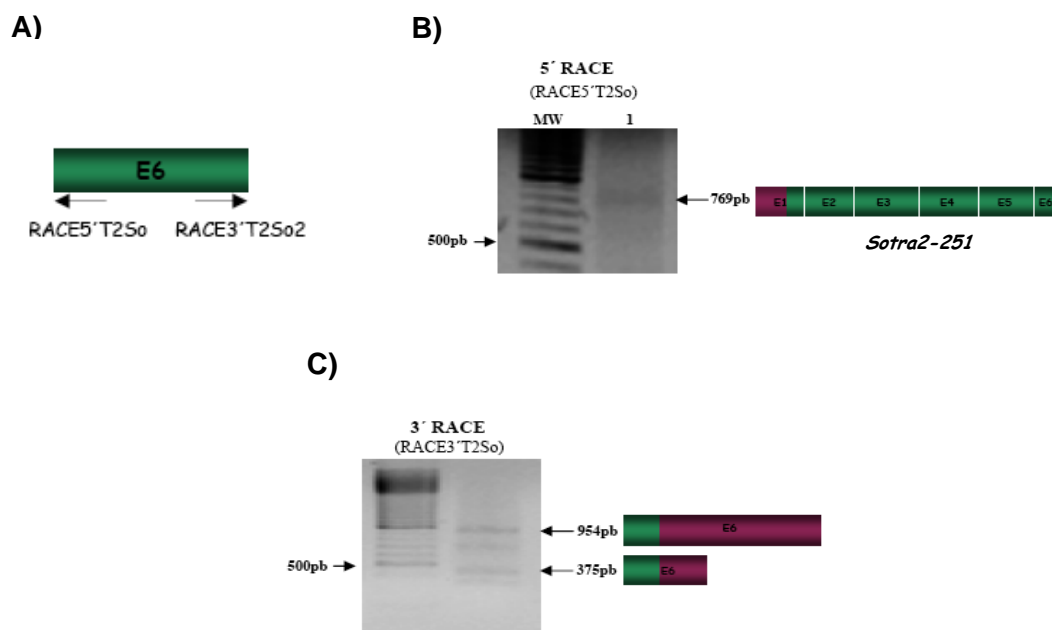
## **4. RESULTADOS**

#### 4.1 AISLAMIENTO, ORGANIZACIÓN MOLECULAR Y TRANSCRITOS DEL GEN *tra2* DE *S.ocellaris*

Con el fin de aislar el gen *tra2* de *S. ocellaris* (*Sotra2*) y determinar su organización molecular se llevaron a cabo diferentes estrategias:

Primeramente se realizó una RT-PCR a partir de ARN total de hembras adultas, utilizando como cebador para la retrotranscripción un *oligo-dT*. Como cebadores de la PCR se utilizaron dos parejas de oligonucleótidos sintéticos degenerados sintetizados a partir de las secuencias pertenecientes a las regiones conservadas del dominio RRM del gen de *tra2* de los dípteros *M. domestica*, *D. virilis* y *D. melanogaster*. Tras la reamplificación, se obtuvo un fragmento de ADNc de 133pb que fue clonado y secuenciado y cuya traducción resultó en una secuencia putativa de aminoácidos que mostraba una similitud del 86% con el dominio RRM del gen *tra2* de *D. melanogaster*. Este resultado indicaba que se había clonado un fragmento de la región carboxilo terminal del gen *tra2* de *S. ocellaris*.

Segundo, con el fin de amplificar la unidad de transcripción completa del gen *Sotra2*, se sintetizaron oligonucleótidos específicos en dirección 5' y 3' del fragmento amplificado que se utilizaron en experimentos 5'RACE y 3'RACE, respectivamente, sobre una genoteca construida a partir de ARN poliadenilado de hembras (Figura 9). En el ensayo 5'RACE (Figura 9B) se amplificó un único fragmento de 769pb, mientras que en el ensayo 3'RACE se amplificaron varios fragmentos (Figura 9C), destacando uno de 375pb y otro de 954pb. Todos los fragmentos amplificados se clonaron y secuenciaron y las secuencias putativas de aminoácidos se compararon con las de las proteínas Tra2 de *Drosophila*. El fragmento amplificado en el 5' RACE contenía la región 5'UTR, el extremo amino terminal, el dominio RS1 y parte del dominio RRM de la proteína putativa SoTra2. Los dos fragmentos amplificados de 375 y 954pb en el 3' RACE contenían el dominio RS2, la región carboxilo terminal de la proteína SoTra2 y la región 3'UTR. La diferencia de tamaño entre estos dos fragmentos amplificados se debe al distinto número de bases que componen la región 3'UTR. Las bandas adicionales que aparecen en el 3' RACE no se correspondían con secuencias de *tra2*.



**Figura 9. RACE 5' y 3'.** (A) Localización sobre el fragmento amplificado de 133pb del gen *Sotra2* de los cebadores "RACE5' T2So" y "RACE3' T2So". (B) RACE 5' con el cebador "RACE5' T2So". (C) RACE 3' con el cebador "RACE3' T2So". Los rectángulos en verde representan los exones. Las regiones 5'UTR y 3'UTR aparecen representadas en morado. Junto a los geles se indica el esquema de los fragmentos amplificados y su tamaño. MW hace referencia al marcador de peso molecular.

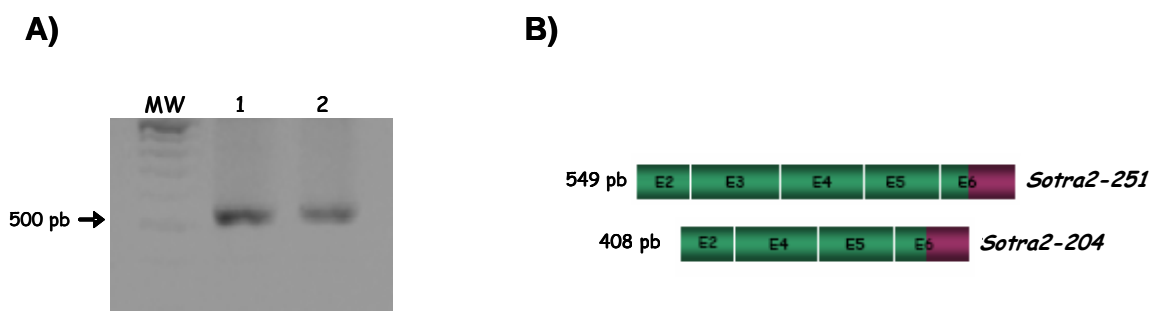
Tercero, se sintetizaron cebadores específicos en dirección 5' y 3' del fragmento amplificado en la RT-PCR, que se utilizaron en ensayos de PCR sobre una genoteca genómica (sintetizada como se describe en el apartado 3.2.2.3 de Materiales y Métodos) utilizando la técnica del paseo cromosómico (*Genome Walker*). Los fragmentos genómicos amplificados se clonaron y secuenciaron, repitiéndose éste proceso hasta conseguir la secuencia genómica completa de 8197pb del gen *tra2* de *S. ocellaris*. La secuencia y localización de los cebadores utilizados se muestran en la Tabla 1 y Figura 8 de Materiales y Métodos. La comparación de las secuencias genómicas con las secuencias de ADNc obtenidas previamente, permitió determinar las uniones exón-intrón, obteniéndose así la organización molecular del gen *tra2* de *S. ocellaris*.

La comparación de la secuencia genómica correspondiente a la región 5' del gen *Sotra2* con la secuencia de ADNc permitió determinar la longitud de la región 5'UTR. Sin embargo, debido a la ausencia de cajas TATA, no fue posible identificar la región promotora del gen *Sotra2*. No obstante, en base a las secuencias analizadas del experimento de 5' RACE, y el programa "Neural Network Promoter Prediction" ([www.fruitfly.org/seq\\_tools/promoter.html](http://www.fruitfly.org/seq_tools/promoter.html)), podemos

determinar que el inicio putativo de la transcripción se encuentra localizado 94 pb aguas arriba del codón de inicio de la traducción.

Los resultados obtenidos hasta el momento indicaban que habíamos aislado un transcrito del gen *Sotra2* de 1177 pb (*Sotra2-251*), formado por seis exones, el cual se correspondía con el ARNm de *tra2-264* de *D. melanogaster*.

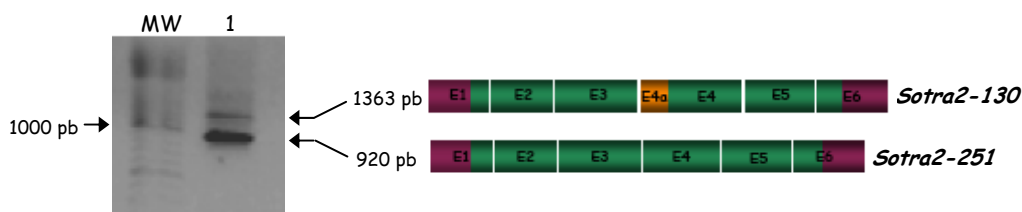
Como se describe en la Introducción, el gen *tra2* de *D. melanogaster* da lugar a varios transcritos distintos por un procesamiento alternativo, que codifican tres isoformas distintas de la proteína Tra2 (Tra2-264, Tra2-226 y Tra2-179) (Amrein *et al.*, 1990; Mattox *et al.*, 1990). Para comprobar si en *S. ocellaris* se producen otros transcritos por un procesamiento alternativo del transcrito primario, se llevaron a cabo ensayos de RT-PCR sobre ARN total de machos y hembras adultos, separadamente, de *S. ocellaris*. Para la retrotranscripción se utilizó un oligo-*dT* y, para la posterior amplificación por PCR, dos cebadores específicos, "*M1So.1*" presente en el exón 2 y "*expSot2.2*" presente en la región 3' UTR en el exón 6. Se amplificó en ambos sexos el fragmento esperado de 549pb, correspondiente al transcrito *Sotra2-251* (Figura 10). El clonaje del producto de esta RT-PCR, permitió identificar otro fragmento de 408pb, no visible en el gel, cuya secuenciación reveló que correspondería a un transcrito de *Sotra2* que carecería del exón 3 (*Sotra2-204*), el cual no sería muy abundante.



**Figura 10. Transcritos de *tra2* en *Sciara ocellaris*.** RT-PCR sobre ARN total de hembras (carril 1) y machos (carril 2) adultos con los cebadores "*Mso.1*" y "*expSot2.2*". Los rectángulos en verde representan los exones. La región 3' UTR aparece representada en morado. Junto a los geles se indica el esquema de los fragmentos amplificados y su tamaño. MW hace referencia al marcador de peso molecular.

En la línea germinal de los machos de *D. melanogaster*, además del transcrito *tra2-226* que codifica la proteína Tra2 funcional, se produce un transcrito específico de la línea germinal, *tra2-179*, que se caracteriza por no procesar el intrón M1 (ver Figura 4 en la Introducción), de manera que el ARNm maduro que contiene en su secuencia dicho intrón, produce una proteína Tra2 truncada no

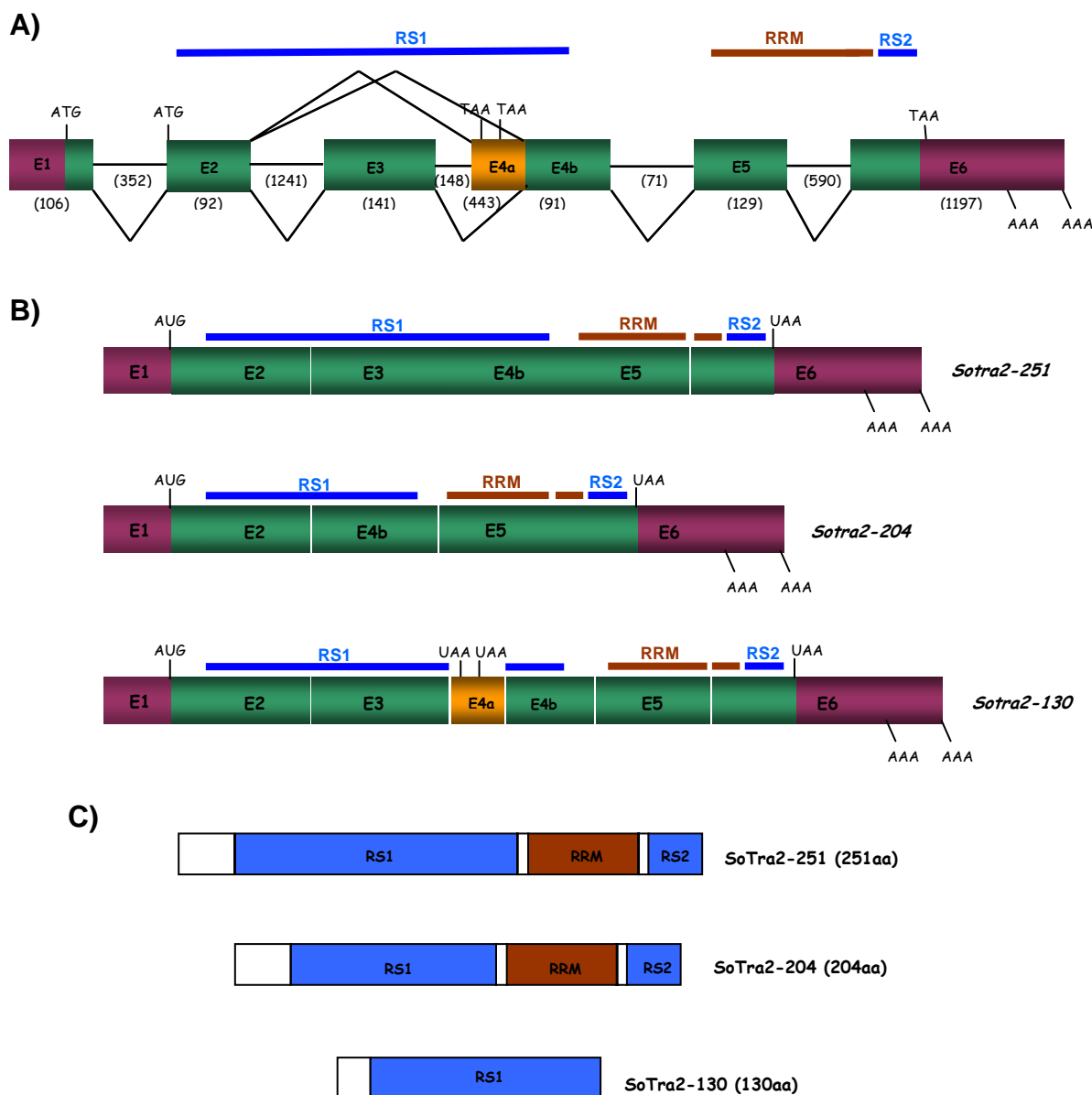
funcional. Con el fin de comprobar si en la línea germinal de los machos de *S. ocellaris* se producía un transcrito similar al transcrito *tra2-179* de *Drosophila*, se llevó a cabo una RT-PCR sobre ARN total de testículos, utilizando para la retrotranscripción un cebador *oligo-dT* y para la amplificación por PCR los cebadores “5´*SoT2*” presente en la región 5´ UTR y “3´*SoT2*” presente en la región 3´ UTR. Se amplificaron dos fragmentos de 920 y 1363 pb, respectivamente, que fueron clonados y secuenciados (Figura 11). El fragmento de 920pb se correspondía con el transcrito *Sotra2-251* mientras que el fragmento de 1363 pb contenía, además, parte de la secuencia correspondiente al intrón 3 (exón E4a). La incorporación de esta parte de dicho intrón en el transcrito maduro, introduce codones de parada de la traducción que da lugar a una proteína putativa SoTra2 truncada y no funcional. Este resultado indicaba que, al igual que en *D. melanogaster*, en *S. ocellaris* existía un transcrito específico de la línea germinal de los machos (*Sotra2-130*). Sin embargo, como se muestra más adelante, este transcrito también se encuentra, aunque en distinta proporción, en las hembras.



**Figura 11. RT-PCR sobre ARN de testículos de *Sciara ocellaris*.** PCR con los cebadores “5´*SoT2*” y “3´*SoT2*”. Los rectángulos en verde representan los exones. Las regiones 5´UTR y 3´UTR aparecen representadas en morado. El intrón que se retiene de forma parcial se muestra en naranja (E4a). Junto a los geles se indica el esquema de los fragmentos amplificados y su tamaño. MW hace referencia al marcador de peso molecular.

Todos estos resultados permitieron determinar la organización molecular del gen *Sotra2* y los distintos tipos de procesamiento de su transcrito primario, que de forma esquemática se presentan en la Figura 12. Este gen está formado por seis exones y cinco intrones, y da lugar, por un procesamiento alternativo, a tres transcritos distintos presentes en ambos sexos: un transcrito de 1177 pb (*Sotra2-251*), formado por los exones E1, E2, E3, E4b, E5 y E6; un transcrito de 1036 pb (*Sotra2-204*) formado por los exones E1, E2, E4b, E5 y E6; y un transcrito de 1620 pb (*Sotra2-130*), formado por los exones E1, E2, E3, E4a, E4b, E5 y E6. El transcrito más abundante es el *Sotra2-251*, el cual codifica una proteína putativa

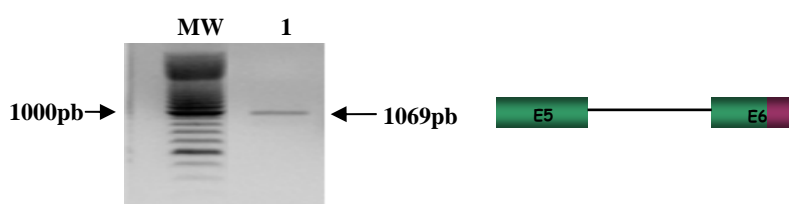
Tra2 que es la homóloga de la proteína Tra2 de *D. melanogaster* y de los otros insectos (ver más adelante).



**Figura 12. Organización molecular y transcritos del gen *tra2* de *S. ocellaris*.** (A) Los rectángulos coloreados en verde representan los exones (E). En naranja se indica el exón E4a. Las regiones 5'UTR y 3'UTR se representan en color morado. Se indican los codones de inicio y de parada de la traducción, los sitios de poliadenilación (AAA) y los dominios RS1, RS2 (líneas azules) y RRM (línea marrón). Las líneas horizontales de color negro representan los intrones. El tamaño de los exones y de los intrones se indica entre paréntesis. El procesamiento alternativo del pre-ARNm se indica mediante líneas de color negro. (B) Transcritos identificados. (C) Esquemas de las proteínas Tra2. Se indica el tamaño de las proteínas en aminoácidos.

## 4.2 AISLAMIENTO, ORGANIZACIÓN MOLECULAR Y TRANSCRITOS DEL GEN *tra2* DE *S.coprophila*

Para llevar a cabo el aislamiento del gen *tra2* de *S. coprophila* se realizó una PCR sobre ADN genómico de machos y de hembras de esta especie. Como cebadores se utilizaron dos oligonucleótidos "*expSoT2.1*" y "*expSoT2.2*", correspondientes a los exones E5 y E6, respectivamente, del gen *Sotra2*, dado que todos los genes clonados en las dos especies de *Sciara* muestran un alto grado de homología a nivel de secuencia nucleotídica. Se amplificó un fragmento de 1069 pb que fue clonado y secuenciado, mostrando su secuencia una identidad del 61% con la secuencia genómica del gen *Sotra2*, lo que indicaba que habíamos clonado un fragmento genómico del extremo carboxilo terminal del gen *tra2* de *S. coprophila* (*Sctra2*) (Figura 13).



**Figura 13. PCR sobre ADN genómico de *Sciara coprophila*.** Se utilizaron los cebadores "*expSoT2.2*" y "*expSoT2.1*". Se indica el tamaño del fragmento amplificado así como su esquema. En verde se representan los exones y en morado la región 3'UTR. MW hace referencia a marcador de peso molecular.

Con el fin de caracterizar la unidad de completa transcripción del gen *Sctra2*, se siguió la misma estrategia que la utilizada para caracterizar el gen *Sotra2*, descrita anteriormente. Primero, se sintetizaron oligonucleótidos específicos de las regiones 5' y 3' del fragmento genómico amplificado que se utilizaron como cebadores en experimentos de 5'RACE y 3'RACE, respectivamente, sobre una genoteca construida a partir de ARN poliadenilado de hembras y machos adultos.

Los fragmentos amplificados se clonaron y secuenciaron, y las secuencias putativas de aminoácidos se compararon con las secuencias de las proteínas Tra2 de *S. ocellaris*. En el ensayo 5'RACE se amplificó, en ambos sexos, un único

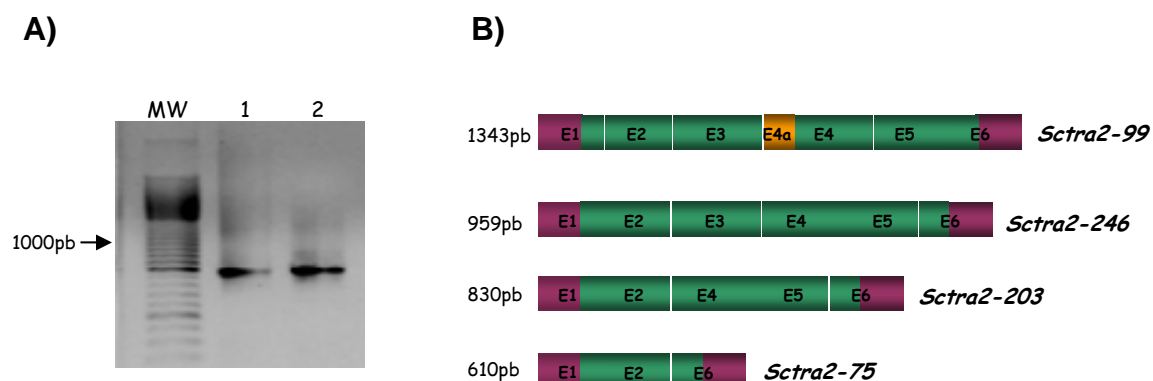


fragmento de 1244 pb que contenía la región 5'UTR, el extremo amino terminal, el dominio RS1 y parte del dominio RRM de la proteína putativa ScTra2. En el ensayo de 3'RACE se amplificó, en ambos sexos, un único fragmento de 799 pb que incluía el dominio RRM, el dominio RS2, la región carboxilo terminal de la proteína ScTra2 y la región 3'UTR.

Segundo, se llevaron a cabo ensayos de PCR sobre una genoteca genómica (sintetizada como se describe en el apartado 3.2.3.3. de Materiales y Métodos) utilizando la técnica del paseo cromosómico (*Genome Walker*). La secuencia y localización de los cebadores utilizados se muestran en la Tabla 1 y Figura 8 de Materiales y Métodos. Los fragmentos genómicos amplificados se clonaron y secuenciaron, repitiéndose éste proceso hasta conseguir la secuencia genómica completa de 5321 pb del gen *tra2* de *S. coprophila*.

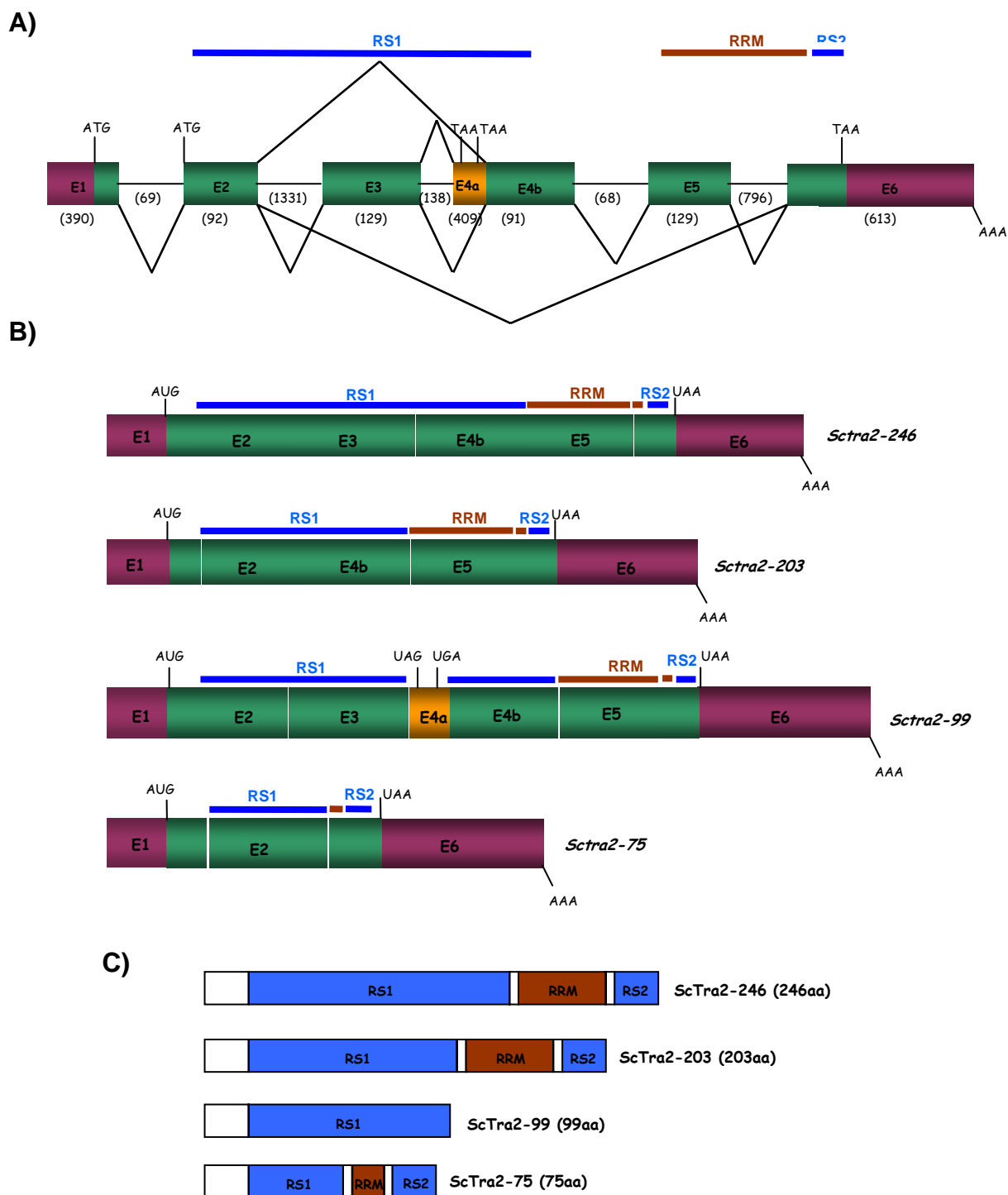
La comparación de las secuencias genómica y del ADNc permitió determinar la organización molecular del gen *tra2* de *S. coprophila*.

Para comprobar si, como en el caso de *S. ocellaris*, en *S. coprophila* existían también varios transcritos alternativos, se llevó a cabo una RT-PCR a partir de ARN total de hembras adultas y testículos de machos adultos utilizando para la retrotranscripción un cebador *oligo-dT* y para la PCR los cebadores "*Scop5´*" y "*Scop3´*" que amplificaban todo el marco abierto de lectura de *ScTra2*. En ambos casos se amplificó un fragmento de 959 pb correspondiente al transcrito que codifica la proteína Tra2 funcional (Figura 14). No obstante, la clonación de los productos de la RT-PCR puso de manifiesto la existencia de otros fragmentos amplificados, difícilmente visibles en el gel, que corresponderían a distintas isoformas del transcrito *tra2*: un fragmento de 1343pb, que se correspondería con el transcrito que retiene parte del intrón 3 (exón E4a); y un fragmento de 830 pb que carece del exón 3. Además, en los testículos se identificó un fragmento adicional de 610 pb, que no se encontraba en las hembras, que correspondería a un transcrito que carece de los exones E3, E4 y E5.



**Figura 14. RT-PCR sobre ARN de *Sciara coprophila*.** RT-PCR sobre ARN total de hembras (carril 1) y de testículos de machos adultos (carril 2) adultos con los cebadores “*Scop5*” y “*S.cop3*”. Se indica el tamaño de los fragmentos amplificados así como su esquema. En verde se representan los exones y en morado la región UTR. MW hace referencia a marcador de peso molecular.

Todos estos resultados permitieron determinar la organización molecular del gen *Sctra2* y los distintos tipos de procesamiento de su transcrito primario, que de forma esquemática se presentan en la Figura 15. El gen *Sctra2* está formado por seis exones y cinco intrones y da lugar a cuatro transcritos distintos por un procesamiento alternativo. Tres de ellos se encuentran tanto en machos como en hembras: un transcrito de 1444 pb (*Sctra2-246*) formado por los exones E1, E2, E3, E4b, E5 y E6; un transcrito de 1315 pb (*Sctra2-203*), que contiene los exones E1, E2, E4b, E5 y E6; y un transcrito de 1853 pb (*Sctra2-99*) formado por los exones E1, E2, E4a, E4b, E5 y E6. El cuarto transcrito de 1095 pb (*Sctra2-75*) formado por los exones E1, E2 y E6, se encuentra únicamente en los testículos.



**Figura 15. Organización molecular y transcritos del gen *tra2* de *S. coprophila*.** (A) Los rectángulos coloreados en verde representan los exones (E). En naranja se indica la parte del intrón 3 que se retiene. Las regiones 5' UTR y 3' UTR se representan en color morado. Se indican los codones de inicio y de parada de la traducción, los sitios de poliadenilación (AAA) y los dominios RS1, RS2 (líneas azules) y RRM (línea marrón). Las líneas horizontales de color negro representan los intrones. El tamaño de los exones y de los intrones se indica entre paréntesis. El procesamiento alternativo del pre-ARNm se indica mediante líneas de color negro. (B) Transcritos identificados. (C) Esquemas de las proteínas ScTra2. Se indica el tamaño de las proteínas en aminoácidos.

### 4.3 COMPARACIÓN DE LA ORGANIZACIÓN MOLECULAR DEL GEN *tra2* DE *Sciara ocellaris* Y *Sciara coprophila* CON EL GEN *tra2* DE OTROS INSECTOS.

La comparación de la organización molecular del gen *tra2* de *Sciara* con el gen *tra2* de otros insectos se muestra en la Figura 16. No se incluye *L. cuprina* pues se desconoce su organización molecular.

El número de exones varía: el gen *tra2* de *Sciara* está formado por 6 exones; el de *Drosophila* por 7 exones; el de *Ceratitis* y *Musca* por 8 exones (teniendo en cuenta 6a y 6b) y el de *Bombyx* por 9 exones (teniendo en cuenta 8a y 8b). En el caso de *Sciara* existe un transcrito que contiene una parte de un intrón (exón E4a), mientras que en el caso de *Drosophila*, existe un transcrito que contiene un intrón (M1) completo. Esto no ocurre en las otras especies estudiadas.

Todas las uniones exón/intrón son las consenso; esto es, GT/AG. Estas uniones están localizadas en el mismo sitio en el caso de *S. ocellaris* y *S. coprophila*, al igual que la unión correspondiente a los exones E5 y E6 que se localiza en el mismo sitio en todas las especies analizadas; las demás uniones exón/intrón varían entre las especies (Ver Figura 17).

El dominio RS1 está codificado por los exones E1-E4, y el dominio RRM está codificado por los exones E5-E6 en todos los casos. El dominio RS2 está codificado por los exones E5-E6 en *Sciara*, E6-E7 en *Drosophila*, E6b-E7 en *Ceratitis*, E6-E7 en *Musca* y E7-E8 en *Bombyx*.

El gen *tra2* de *Sciara* posee dos marcos de lectura abiertos putativos, uno en el exón E1 y el otro en el exón E2, que están en fase; de modo que ambos producen la misma proteína excepto por la adición de cuatro amino ácidos en el extremo amino-terminal en el caso del marco abierto de lectura mayor.

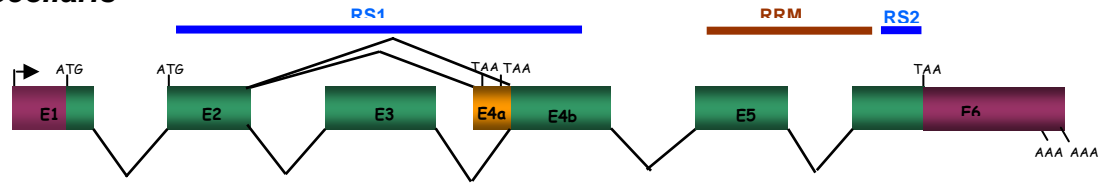
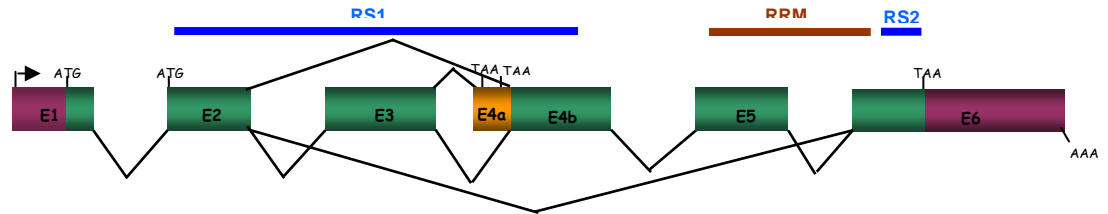
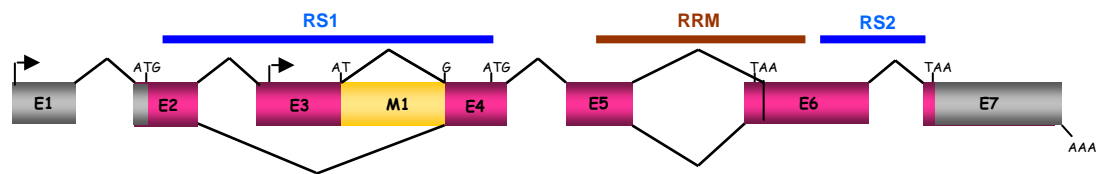
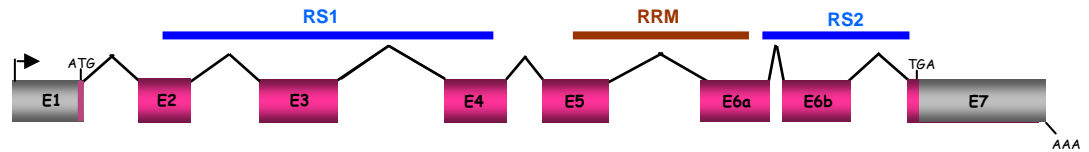
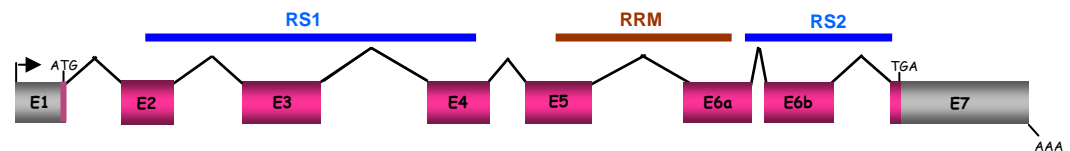
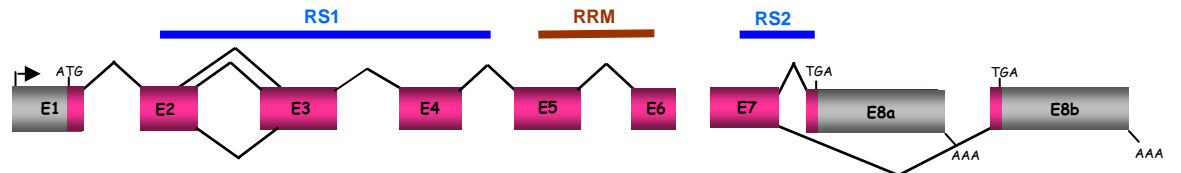
**A) *S. ocellaris*****B) *S. coprophila*****C) *D. melanogaster*****D) *C. capitata*****E) *M. domestica*****F) *B. mori***

Figura 16. de la organización molecular de los genes *tra2* de *S. ocellaris* (A), *S. coprophila* (B), *D. melanogaster* (C), *C. capitata* (D) *M. domestica* (E) y *B. mori* (F). Los rectángulos coloreados representan los exones. Se indican los codones de inicio y de parada de la traducción, los sitios de poliadenilación (AAA) y los dominios RS1, RS2 (líneas azules) y RRM (línea marrón).

#### 4.4 LA PROTEÍNA TRA2 DE *S. ocellaris* Y *S. coprophila* Y SU COMPARACIÓN CON LAS PROTEÍNAS TRA2 DE OTROS INSECTOS.

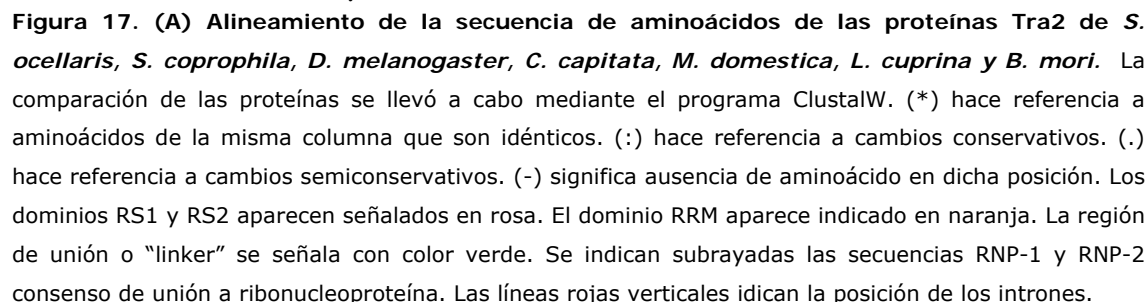
En *S. ocellaris*, el transcrito *Sotra2-251* daría lugar a una proteína putativa de 251 aminoácidos (SoTra2-251) que contiene el dominio RS1 completo, el dominio RRM y el dominio RS2. El transcrito *Sotra2-204* daría lugar a una proteína putativa de 204 aminoácidos (SoTra2-204) con el dominio RS1 más corto, debido al procesamiento del exón 3, el dominio RRM y el dominio RS2. El transcrito *Sotra2-130* daría lugar a una proteína putativa de 130 aminoácidos (SoTra2-130) conteniendo solamente el extremo amino-terminal y parte del dominio RS1, careciendo de los dominios RRM y RS2, debido a la incorporación del exón E4a, que introduce codones de parada de la traducción.

En *S. coprophila*, el transcrito *Sctra2-246* daría lugar a una proteína putativa de 246 aminoácidos (ScTra2-246) que contiene el dominio RS1 completo, el dominio RRM y el dominio RS2. El transcrito *Sctra2-203* daría lugar a una proteína putativa de 203 aminoácidos (ScTra2-203) con un dominio RS1 más corto, debido al procesamiento del exón 3, y con el dominio RRM y el dominio RS2. El transcrito *Sctra2-99* daría lugar a una proteína putativa de 99 aminoácidos (ScopTra2-99) conteniendo solamente el extremo amino-terminal y parte del dominio RS1, faltándole los dominios RRM y RS2, debido a la incorporación del exón E4a, que introduce codones de parada de la traducción. Finalmente, el transcrito *Sctra2-75* daría lugar a una proteína putativa de 75 aminoácidos (ScTra2-75) conteniendo los dominios RS1 y RRM truncados y el dominio RS2 completo.

El alineamiento de la secuencia de las proteínas putativas SoTra2 y ScTra2 con las proteínas Tra2 de otros insectos se muestra en la Figura 17. En los casos en los que *tra2* ha sido caracterizado en varias especies de un mismo género, se ha utilizado una especie de dicho género como referencia. Además, en aquellas especies en las que existe más de una isoforma de la proteína Tra2 se ha elegido la isoforma que presenta la proteína completa. Así la comparación se ha realizado entre las especies de dípteros pertenecientes al suborden Nematocera *Sciara ocellaris* y *Sciara coprophila* (familia Sciaridae), y dípteros pertenecientes al suborden Brachycera *Drosophila melanogaster* (familia Drosophilidae), *Ceratitis capitata* (familia Tephritidae), *Musca domestica* (familia Muscidae), *Lucilia cuprina* (familia Calliphoridae) así como con el lepidóptero *Bombyx mori* (suborden Glossata, familia Bombycidae). La comparación de las proteínas se llevó a cabo a

nivel de las secuencias completas, así como a nivel de las secuencias específicas de los distintos dominios (Tablas 2A-F).

El número de aminoácidos ("aa" entre paréntesis) que componen estas proteínas varía entre las distintas especies, existiendo la mayor similitud entre las proteínas Tra2 de *S. ocellaris* y *S. coprophila* (88,6%), como era de esperar al tratarse de dos especies de un mismo género, luego con *Musca* (46,9%) y *Bombyx* (44,6%), después con *Ceratitis* (42,6%), y con *Lucilia* (39,4%), y finalmente con *Drosophila* (36,3%) (Tabla 2A). El mayor grado de similitud se observa en el dominio RRM (entre 55 y 94%) (Tabla 2B), y en la región "linker" (entre 68 y 100%) (Tabla 2C), que une el RRM con el dominio RS2, conservándose, en todas ellas, el número de aminoácidos que componen dicho dominio RRM y la región "linker". La similitud en los dominios RS1 (Tabla 2D) y RS2 (Tabla 2E) es menor, variando el número de aminoácidos que los componen. Sin embargo, todos ellos contienen un número abundante, aunque variable, de dipéptidos arginina-serina (Tabla 2F).





**Tabla 2A.** Porcentaje de similaridad (identidad de aminoácidos y cambios conservativos) de las proteínas Tra2.

	<i>S. coprophila</i> (245 aa)	<i>D. melanogaster</i> (264 aa)	<i>C. capitata</i> (251 aa)	<i>M. domestica</i> (232 aa)	<i>L. cuprina</i> (271 aa)	<i>B. mori</i> (274 aa)
<i>S. ocellaris</i> (251 aa)	88,6	36,3	42,6	46,9	39,4	44,6

**Tabla 2B.** Porcentaje de similaridad (identidad de aminoácidos y cambios conservativos) del dominio RRM de las proteínas Tra2.

	<i>S. coprophila</i> (72 aa)	<i>D. melanogaster</i> (72 aa)	<i>C. capitata</i> (72 aa)	<i>M. domestica</i> (72 aa)	<i>L. cuprina</i> (72 aa)	<i>B. mori</i> (72 aa)
<i>S. ocellaris</i> (72 aa)	94,4	55,5	68,1	72,2	65,3	72,2

**Tabla 2C.** Porcentaje de similaridad (identidad de aminoácidos y cambios conservativos) entre la región "linker" de las proteínas Tra2.

	<i>S. coprophila</i> (19 aa)	<i>D. melanogaster</i> (19 aa)	<i>C. capitata</i> (19 aa)	<i>M. domestica</i> (19 aa)	<i>L. cuprina</i> (19 aa)	<i>B. mori</i> (19 aa)
<i>S. ocellaris</i> (19 aa)	100	68,4	73,7	73,7	73,7	73,7

**Tabla 2D.** Porcentaje de similaridad (identidad de aminoácidos y cambios conservativos) del dominio RS1 de las proteínas Tra2.

	<i>S. coprophila</i> (95 aa)	<i>D. melanogaster</i> (65 aa)	<i>C. capitata</i> (90 aa)	<i>M. domestica</i> (87 aa)	<i>L. cuprina</i> (109 aa)	<i>B. mori</i> (24 aa)
<i>S. ocellaris</i> (99 aa)	87,6	23,8	28,7	28,7	21,8	28,7

**Tabla 2E.** Porcentaje de similaridad (identidad de aminoácidos y cambios conservativos) del dominio RS2 de las proteínas Tra2.

	<i>S. coprophila</i> (38 aa)	<i>D. melanogaster</i> (67 aa)	<i>C. capitata</i> (51 aa)	<i>M. domestica</i> (42 aa)	<i>L. cuprina</i> (67 aa)	<i>B. mori</i> (35 aa)
<i>S. ocellaris</i> (40 aa)	81,6	38,5	35,9	35,9	28,2	41,0

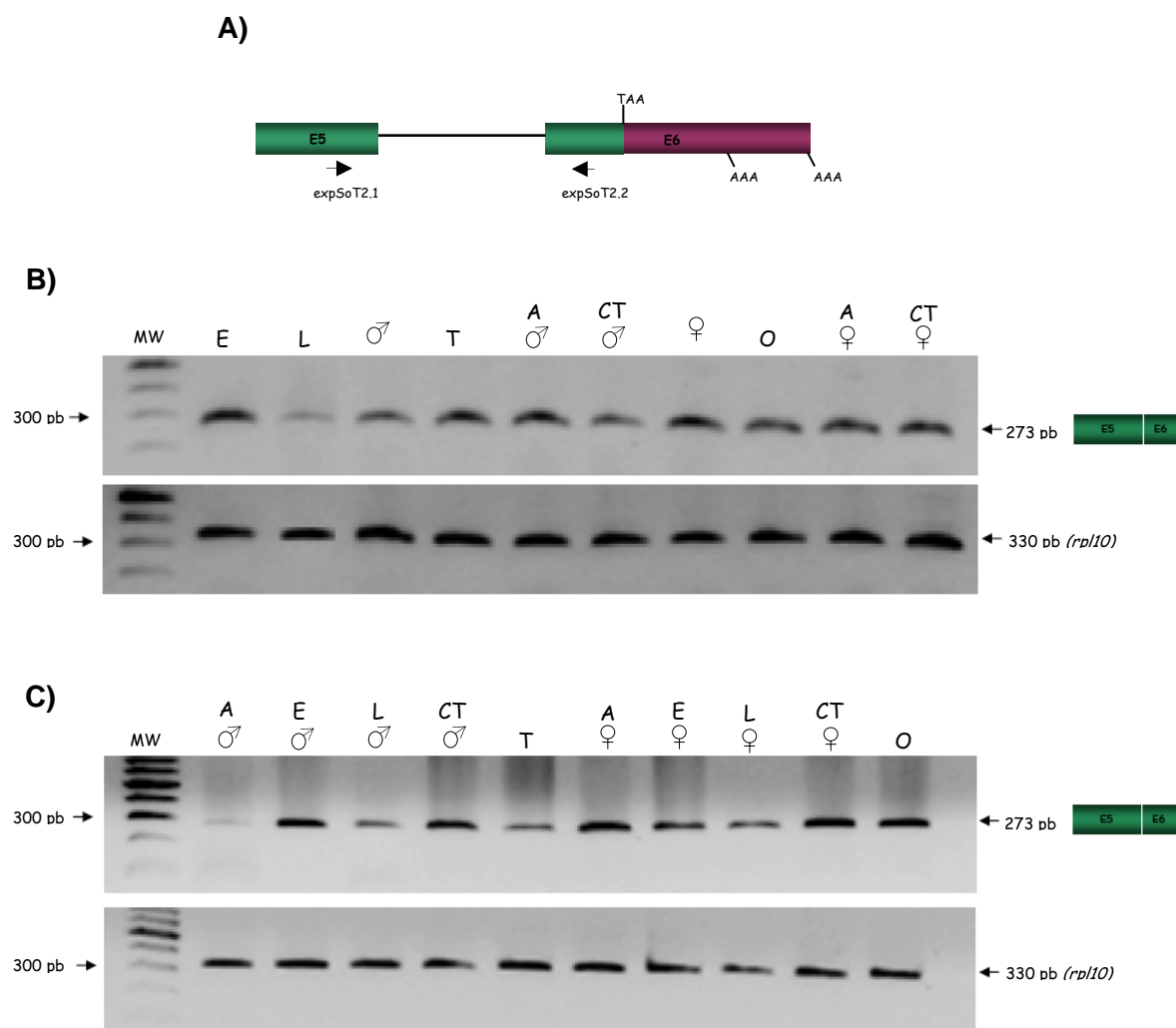
**Tabla 2F.** Número de dipéptidos arginina-serina en los dominios RS1 y RS2 de las proteínas Tra2.

	<i>S. ocellaris</i>	<i>S. coprophila</i>	<i>D. melanogaster</i>	<i>C. capitata</i>	<i>M. domestica</i>	<i>L. cuprina</i>	<i>B. mori</i>
<b>Dominio RS1</b>	18	15	11	12	12	16	16
<b>Dominio RS2</b>	4	5	6	4	5	3	4

#### 4.5 PATRÓN DE EXPRESIÓN DEL GEN *tra2* DE *S. ocellaris* Y *S. coprophila*

En aquellas especies donde *tra2* ha sido caracterizado, éste se expresa durante todo el desarrollo y en todos los tejidos de los adultos. Con el fin de determinar el patrón de expresión de *tra2* en *S. ocellaris* y en *S. coprophila*, se llevaron a cabo análisis de RT-PCR sobre ARN total de embriones hembras y machos de 24 horas, larvas hembras y larvas machos, hembras adultas, machos adultos, cabeza y tórax de hembras, cabeza y tórax de machos, abdómenes de hembras, abdómenes de machos, ovarios y testículos de adultos. En el caso de *S. ocellaris*, los embriones y las larvas correspondían a una mezcla de ambos sexos, mientras que en el caso de *S. coprophila*, se analizaron separadamente tanto los embriones como las larvas de cada sexo. La retrotranscripción se realizó utilizando un *oligo-dT*. Los cebadores utilizados en la reacción de PCR, tanto para *S. ocellaris* como para *S. coprophila*, fueron "*expSoT2.1*" y "*expSoT2.2*" localizados en los exones 5 y 6, respectivamente, (Figura 18A) dado que sus secuencias son idénticas en ambas especies. Como control se utilizó el gen *rpl10*, que codifica la proteína ribosómica Rpl10.

El gen *tra2* de *S. ocellaris* y de *S. coprophila* se expresa tanto en las hembras como en los machos, durante su desarrollo embrionario y larvario, así como en el estado adulto, incluyendo los testículos y los ovarios, lo que indica que este gen posee expresión materna (Figura 18B), tal como ocurre en otros insectos.



**Figura 18. Expresión de los genes *Sotra2* y *Sctra2* a lo largo del desarrollo en *S. ocellaris* y *S. coprophila*.** (A) Localización en el gen *Sotra2* de los cebadores (flechas) empleados en las reacciones de PCR. Los rectángulos verdes representan los exones. En morado la región 3'UTR. Se indica el codón de parada de la traducción y los sitios de poliadenilación (AAA). La línea negra representa el intrón. (B) Patrón de expresión del gen *Sotra2* en los estadios de embrión de 24h (E), larvas hembras y machos conjuntamente (L), hembras adultas (♀), machos adultos (♂), cabeza y tórax de macho (CT♂), hembra (CT♀), abdomen de macho (Ab♂), hembra (Ab♀), testículos (T) y ovarios (O) adultos. Como control se utilizó el gen que codifica la proteína ribosómica Rpl10. (C) Patrón de expresión del gen *Sctra2* en los estadios de embrión macho (E♂) y hembra (E♀) de 24h, larva macho (L♂) y hembra (L♀), hembras adultas (A♀), machos adultos (A♂), cabeza y tórax de macho (CT♂) y hembra (CT♀), testículos (T) y ovarios (O) adultos. Como control se utilizó el gen que codifica la proteína ribosómica Rpl10. Se indica el tamaño de los fragmentos amplificados. MW hace referencia al marcador de peso molecular.

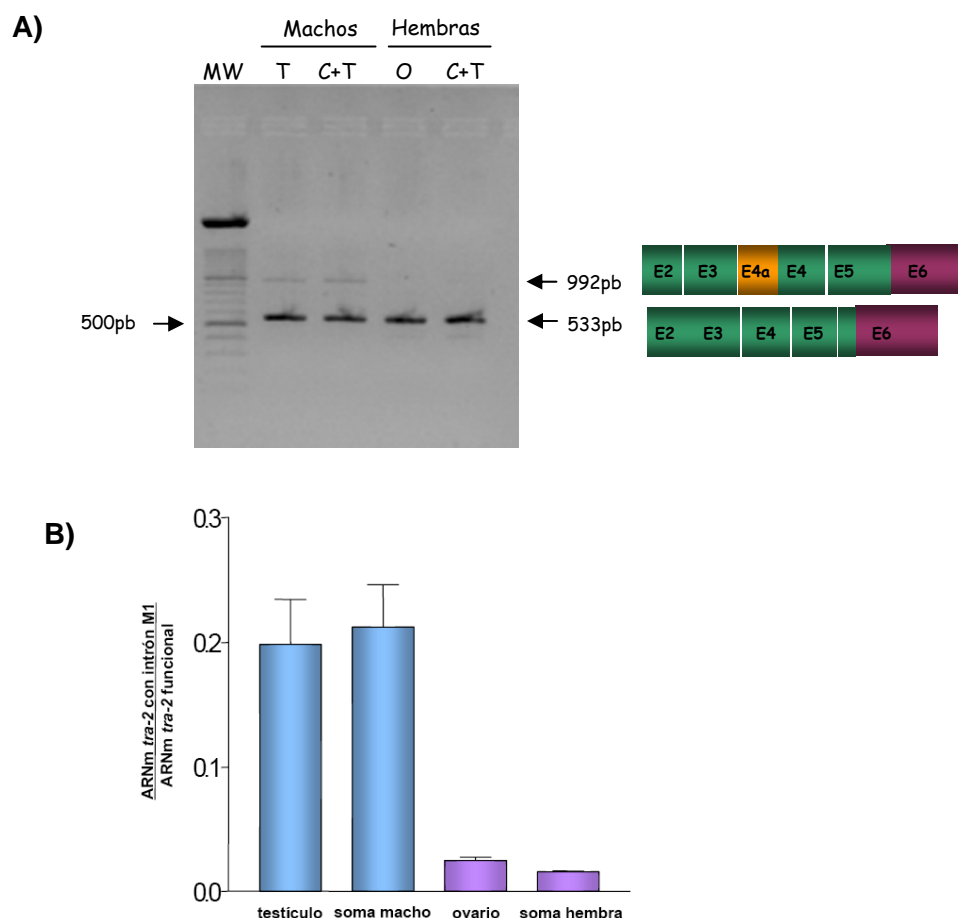
#### 4.6 EL TRANSCRITO *Sotra2-130* ESTÁ PRESENTE TANTO EN MACHOS COMO EN HEMBRAS.

Como se ha mencionado anteriormente, en *Drosophila*, pero no en las otras especies fuera de los drosophilidos donde *tra2* ha sido caracterizado, se producen dos transcritos en la línea germinal de los machos: el transcrito *tra2-226*, que codifica la proteína Tra2 funcional, y el transcrito *tra2-179*, que retiene el intrón M1, codificando una proteína Tra2 truncada. Esta retención está mediada por la interacción de la propia proteína Tra2 con secuencias ISS específicas localizadas en el intrón M1.

En el ensayo de RT-PCR sobre ARN total de testículos de *S. ocellaris*, aislamos un transcrito, *Sotra2-130*, que retiene en su ARNm parte del intrón 3 (exón E4a) que introduce codones de parada de la traducción dando lugar a una proteína Tra2 truncada. Dicho exón contiene dos secuencias ISS, encontradas en el intrón M1 de *tra2* de *D. melanogaster*, en el caso de *S. ocellaris*, y una secuencia en el caso de *S. coprophila*. Esto podría ser una indicación de que el transcrito *Sotra2-130* encontrado en *Sciara* correspondiese al transcrito *tra2-179* específico de la línea germinal del macho de *Drosophila*, y al igual que en esta especie, existiese en *Sciara* una autorregulación negativa del gen *tra2* específica de la línea germinal del macho.

Con el fin de comprobar si el transcrito *Sotra2-130* era específico de la línea germinal de macho de *S. ocellaris*, o si por el contrario, se producía también en las hembras y/o en otros tejidos somáticos, se llevó a cabo una RT-PCR a partir de ARN total de cabeza y tórax de hembras y de macho, por separado, y de testículos y ovarios adultos. Para la retrotranscripción se usó como cebador un *oligo-dT*, y para la amplificación los oligonucleótidos, "*Socel3.1*", situado en el exón E4a, específico de este transcrito, y "*expSoT2.1*" localizado en el exón E6, común a todos los transcritos. Los resultados se muestran en la Figura 19. En todos los casos se amplificó un fragmento de 992 pb del tamaño esperado y cuya secuencia se correspondía con el transcrito *Sotra2-130* (Figura 19A). Por lo tanto, y al igual que ocurre en el resto de los insectos donde *tra2* ha sido caracterizado, excepto en *Drosophila*, este gen en *S. ocellaris* no produce ningún transcrito específico de la línea germinal de los machos.

Con el fin de determinar si los niveles de expresión de este transcrito, con respecto a los otros transcritos identificados, eran similares en machos y en hembras y/o en la línea germinal y somática, se empleó la siguiente estrategia. Se llevaron a cabo ensayos de RT-PCR sobre ARN total de cabeza y tórax de machos y de hembras, por separado, ovarios y testículos adultos, utilizando para la retrotranscripción un cebador *oligo-dT*. Para la amplificación por PCR se utilizaron los cebadores "*M1So.1*" y "*expSo.T2.2*" localizados en los exones E2 y E6, respectivamente. Esta pareja de cebadores permitía amplificar los tres transcritos identificados y así poder medir la cantidad relativa de cada uno de ellos. En primer lugar, para definir las condiciones óptimas que nos permitieran amplificar sin llegar a la saturación, se realizaron PCRs en las que se variaba, únicamente, el número de ciclos de amplificación (Apartado 3.2.2.7 de Materiales y Métodos). Una vez establecidas estas condiciones se realizaron tres réplicas (un ejemplo se muestra en la Figura 19A). Los fragmentos amplificados se cargaron en un gel de electroforesis y se cuantificaron utilizando el programa "Quantity One". En todos los casos se amplificaron dos fragmentos, que fueron clonados y secuenciados. El fragmento de mayor tamaño se correspondía con el transcrito *Sotra2-130* mientras que el fragmento más pequeño correspondía al transcrito *Sotra2-251* que codifica la proteína Tra2 completa. En ninguno de los casos se amplificó el transcrito *Sotra2-204*, lo que parece indicar que este transcrito se encuentra poco representado.



**Figura 19. Análisis de la expresión del transcrito *Sotra2-130* en *S. ocellaris*.** (A) RT-PCR sobre ARN total de testículos (T), cabeza y tórax (C+T) y ovarios (O). Se muestra el esquema de los transcritos correspondientes a cada fragmento amplificado y su tamaño. MW hace referencia al marcador de peso molecular. (B) Comparación de la expresión de los transcritos *Sotra2-251* y *Sotra2-130*.

La abundancia relativa del transcrito *Sotra2-130* (con el exón E4a) con respecto al transcrito *Sotra2-251* (codifica Tra2 funcional) en el soma y la línea germinal de hembras y machos se muestra en la Figura 19B. La cantidad de transcrito *Sotra2-251* no difiere significativamente ( $P > 0,05$ , test ANOVA factor-2) entre hembra y macho ni entre soma y línea germinal (datos no mostrados), mientras que sí existe una diferencia significativa ( $P < 0,001$ ) en la relación *Sotra2-130/Sotra2-251*, siendo más abundante la isoforma *Sotra2-130* en machos que en hembras (test ANOVA factor-2). Dentro de cada sexo, la relación entre los dos transcritos no difiere significativamente ( $P > 0,05$ ) entre el soma y la línea germinal en los machos, pero sí es significativa ( $P > 0,05$ ) en las hembras (t-test). Por lo tanto, aunque la expresión del transcrito *Sotra2-130* no es específica de sexo, sí parece serlo la cantidad del mismo producida en cada uno de ellos.

#### 4.7 EFECTO DEL GEN *tra2* DE *Sciara ocellaris* EN EL DESARROLLO SEXUAL DE *Drosophila melanogaster*.

El análisis consistió en expresar el gen *tra2* de *S. ocellaris*, mediante el sistema GAL4-UAS (descrito en Materiales y Métodos), en pseudomachos de *D. melanogaster*; esto es, en moscas cromosómicamente XX pero que se desarrollan como machos por faltarles la función *tra2* endógena, y monitorizar el fenotipo sexual estudiando las regiones de dimorfismo sexual, tales como el basitarsus del primer par de patas; los tergitos 5, 6 y 7; y la terminalia externa. El basitarsus de los machos se caracteriza por llevar el peine sexual, carente en las hembras. Los tergitos 5 y 6 de los machos están totalmente pigmentados, mientras que los de las hembras poseen pigmentación solamente en su región posterior. Finalmente, la terminalia externa es la región que presenta mayor dimorfismo sexual, estando formada por estructuras de la genitalia, específicas de cada sexo, y de la analia. En el caso de los machos, las estructuras de la genitalia son el arco genital, los clasper, las placas laterales, el hypandrium y el aparato del penis. En el caso de las hembras estas estructuras son el tergito 8 y las placas vaginales. La analia está presente en ambos sexos pero su morfología es totalmente distinta en machos y hembras.

El experimento consistió en cruzar moscas portadoras del trangén *TDSO* con moscas portadoras del factor GAL4 bajo el control de un promotor de expresión generalizada, y generar seudomachos portadores del transgén *TDSO* y del factor GAL4; monitorizando el fenotipo sexual de las regiones de dimorfismo sexual anteriormente citadas; esto es, determinar si la expresión del transgén *TDSO*, inducida por la expresión del GAL4, produce una transformación del desarrollo sexual de macho a hembra. Se utilizaron cuatro líneas transgénicas portadoras del transgén *TDSO*, y dos líneas portadoras del factor GAL4; en una de ellas, el GAL4 estaba bajo el control del promotor del gen *daughterless* (línea *da-GAL4*) y la otra bajo el control del promotor de choque térmico (línea *HS-GAL4*). Como era de esperar, ninguna de las líneas transgénicas de *D. melanogaster* portadoras del transgén *TDSO* (*Sotra2-UAS*) lo expresaban en ausencia del factor GAL4. Las moscas que expresaban el transgén *TDSO* resultaron ser letales para ambas líneas *da-GAL4* y *HS-GAL4*, independientemente de que fuesen seudomachos XX o machos XY.

Puesto que la expresión generalizada de *TDSO* causa letalidad, probamos una estrategia alternativa, consistente en expresar dicho transgén en una región localizada. Específicamente, usamos la línea *rn-GAL4* que determina la expresión de GAL4 en el disco de pata durante la segunda mitad del tercer estadio larvario, expresándose este factor en el mismo dominio que el gen *rotund (rn)*, requerido para el desarrollo de los segmentos tarsales (St. Pierre *et al.*, 2002). Dado que *rotund* se expresa, también, durante la embriogénesis, y que la eficiencia de GAL4 para activar las secuencias UAS es dependiente de la temperatura (menos eficiente a menos temperatura) (Duffy, 2002), realizamos un cruce para generar pseudomachos y sus hermanos XY, ambos mutantes para *tra2*, que portasen el transgén *TDSO* y *rn-GAL4*, los cuales se desarrollaron a 18°C hasta el tercer estadio larvario, y a partir de este momento se desarrollaron a 25°C hasta la eclosión de los adultos. La idea era intentar que la expresión del transgén *TDSO* al principio del desarrollo fuese menor, esperando que no causase letalidad, y que posteriormente a 25°C la expresión fuese mayor pero localizada en la pata, de modo que pudiésemos analizar el fenotipo sexual del basitarsus del primer par de patas. Sin embargo, las moscas que expresaban el transgén seguían siendo letales. Por lo tanto, la expresión del gen *tra2* de *S. ocellaris* causa letalidad en *D. melanogaster*. Dicho resultado no es, en principio, extraño, pues la expresión generalizada de su propio gen *tra2* causa, también, letalidad de *D. melanogaster* (Qi *et al.*, 2006). Dado que la propia proteína Tra2 es un factor del procesamiento, es muy probable que su expresión generalizada interfiera negativamente con la maquinaria del procesamiento.

Puesto que el efecto directo del gen *TDSO* en el desarrollo sexual de *D. melanogaster* no ha podido ser estudiado, por la letalidad que causa su expresión, hemos utilizado una alternativa basada en el hecho de que el desarrollo y la diferenciación de las regiones de dimorfismo sexual que caracterizan a cada sexo depende del gen *dsx*: la proteína DsxF determina desarrollo de hembra y la proteína DsxM determina desarrollo de macho. Como se mencionó en la Introducción, el procesamiento alternativo del pre-ARNm de *dsx*, para dar lugar a DsxF o DsxM, depende de la presencia del complejo Tra-Tra2. Por lo tanto, estudiamos el efecto de *TDSO* en el desarrollo sexual de *D. melanogaster* analizando el patrón de expresión del pre-ARNm de su gen *dsx* endógeno en presencia de proteína transgénica SoTra2. Este análisis se llevó a cabo en pseudomachos de genotipo *yw/w; Df(2R)trix,tra2[-]/tra2<sup>B</sup>; TDSO#/HS-GAL4* (moscas experimentales) y sus hermanos XY de genotipo *yw/Y; Df(2R)trix,tra2[-*



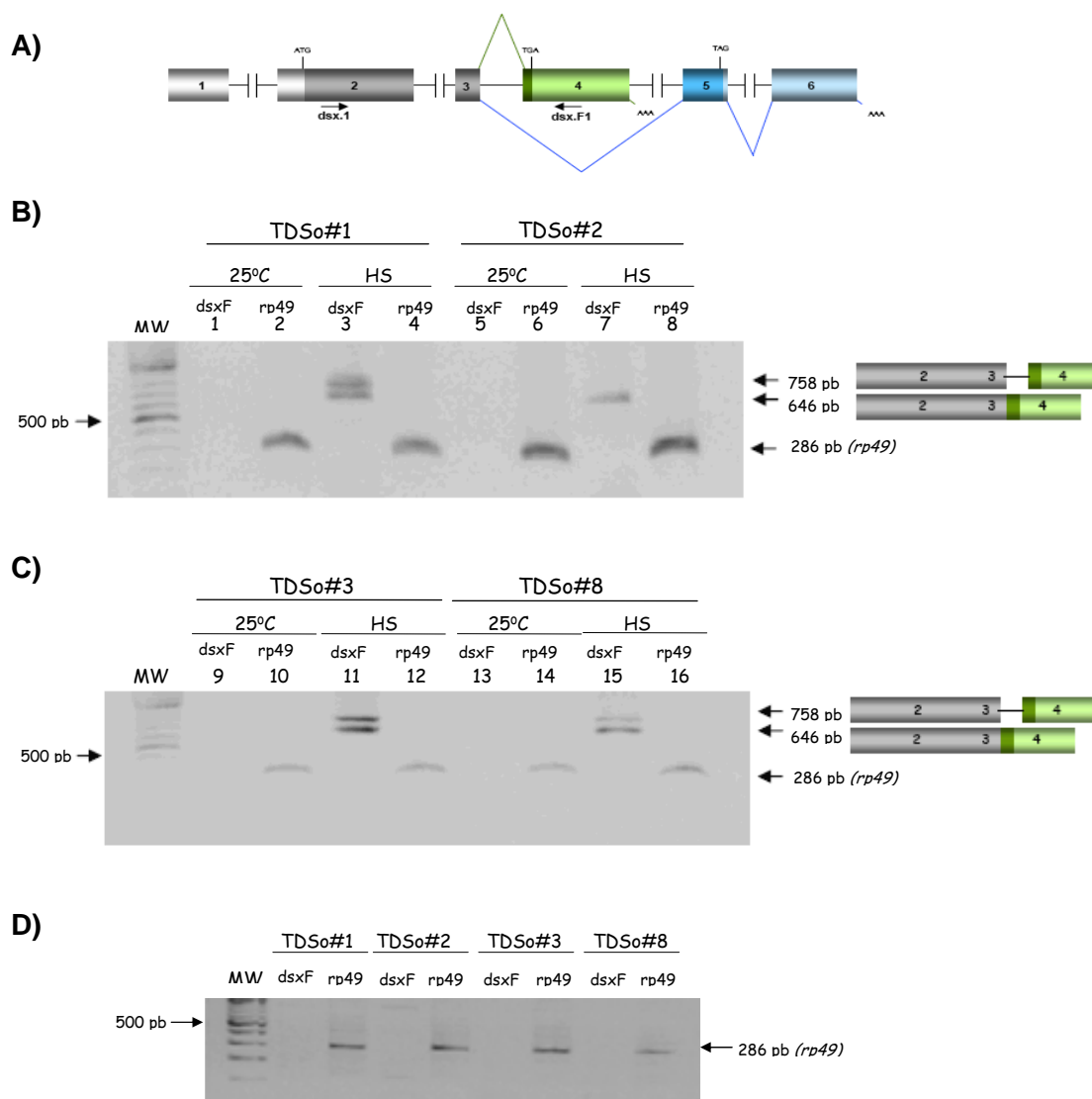
*J/tra2<sup>B</sup>; TDSO#/HS-GAL4* (machos control). Estas moscas se desarrollan a 25°C, alcanzando el estado adulto, pues el transgén *TDSO* no se expresa a esa temperatura. Los adultos de cada genotipo, los cuales expresan el transcrito *dsxM*, se dividieron en dos grupos, uno se mantuvo a 25°C y el otro se sometió a choque térmico para que se expresara el transgén *TDSO*. El efecto del transgén *TDSO* sobre el procesamiento del pre-ARNm *dsx* endógeno se llevó a cabo mediante RT-PCR sobre ARN total. Como control se utilizó la expresión del gen *rp49* que codifica la proteína ribosómica Rp49. Para los detalles experimentales ver el Apartado 3.2.4.6 de Materiales y Métodos. Se estudiaron cuatro líneas transgénicas. Los resultados se muestran en la Figura 20.

En las cuatro líneas transgénicas analizadas, las moscas experimentales no expresaban el transcrito específico de hembra *dsxF* cuando no recibían el choque térmico (Figura 20 B, C, carriles 1, 5, 9 y 13). Sin embargo, sí se detectaba dicho transcrito cuando estas moscas eran sometidas a choque térmico y, por lo tanto, expresaban el transgén *TDSO* (carriles 3, 7, 11 y 15). La cantidad de transcrito *dsxF* varía de unas líneas a otras, lo que puede deberse a diferentes niveles de expresión del transgén, dependiendo de su sitio de inserción en el genoma. En algunos casos (carriles 3, 11 y 15), se observaban dos fragmentos de 654 y 758pb, que se clonaron y secuenciaron. El fragmento de menor tamaño se correspondía con el ARNm de *dsxF* que codifica la proteína normal DsxF, mientras que el fragmento de mayor tamaño se correspondía con una isoforma del ARNm que retiene el intrón 3 entre el exón 3, común a ambos sexos, y el exón 4 específico de hembra.

Los machos control no expresaban el transcrito *dsxF* tras ser sometidos al choque térmico (Figura 20D, carriles 1, 3, 5 y 7), lo que demuestra que primero, la presencia del transcrito *dsxF*, en las moscas experimentales, es consecuencia de la expresión del transgén *TDSO*, y no del choque térmico en sí, y segundo, que el efecto de la proteína transgénica Tra2 de *Sciara* en el desarrollo sexual de *Drosophila* es dependiente de la proteína endógena Tra de esta especie, ya que que los pseudomachos XX poseen proteína Tra, mientras que los machos XY carecen de ella (ver Introducción).

Las moscas experimentales así como las moscas control expresaron el gen *rp49* utilizado como control (Figura 20B, C, D).

Por lo tanto podemos concluir que el gen *tra2* de *S. ocellaris* es capaz de proveer la función *tra2* en *D. melanogaster*.



**Figura 20. Expresión del gen *doublesex* en moscas de *D. melanogaster* transgénicas para *Sotra2*.**

**(A)** Localización en el gen *dsx* de *D. melanogaster* de los cebadores utilizados (flechas negras). Los rectángulos coloreados representan los exones: en gris los exones comunes a ambos sexos, en verde el exón específico de hembra y en azul los exones específicos de macho. Las regiones 5'UTR y 3'UTR se representan un tono más claro. Se indican los codones de inicio y de parada de la traducción así como los sitios de poliadenilación (AAA). Las líneas de color verde y azul representan el procesamiento alternativo específico de hembra y de macho, respectivamente. La RT-PCR se llevó a cabo sobre ARN total de seudomachos (XX) *yw/w*; *Df(2R)trix,tra2[-]/tra2<sup>B</sup>*; *TDSO#1/HS-GAL4* **(B, C)** y de sus hermanos *yw/Y*; *Df(2R)trix,tra2[-]/tra2<sup>B</sup>*; *TDSO#1/HS-GAL4* **(D)** procedentes del cruce de hembras *yw*; *tra2<sup>B</sup>/CyO*; *TDSO#1/MKRS,Sb* con machos *w/Y*; *Df(2R)trix,tra2[-]/CyO*; *HS-GAL4*. *TDSO#* hace referencia a la línea transgénica utilizada. *DsxF* hace referencia a al transcrito *dsxF* específico de hembra. *rp49* hace referencia a la expresión del gen, utilizado como control, que codifica la proteína ribosómica Rp49. C hace referencia a las moscas control que se mantuvieron a 25°C y HS a las moscas que fueron sometidas a choque térmico. Se indica el esquema de los fragmentos amplificados así como su tamaño. MW hace referencia al marcador de peso molecular.

## **5. DISCUSIÓN**

En este trabajo de tesis doctoral se presenta el aislamiento y la caracterización en *S. ocellaris* y *S. coprophila* del gen ortólogo de *tra2* que se ha aislado también en otros insectos. Existen varios datos que nos indican que hemos aislado dicho gen. Primero, cuando se comparan a nivel global las proteínas putativas codificadas por *tra2* de *Sciara* en los bancos de datos encontramos que la mayor similitud se corresponde con las proteínas Tra2 de otros insectos. Segundo, las proteínas Tra2 de *Sciara* poseen los dominios que caracterizan a las proteínas Tra2 de otros insectos, esto es, los dominios RS y el dominio RRM de unión a ARN. Tercero, las proteínas putativas producidas por el gen de *Sciara*, corresponden a ARN mensajeros que se generan por un procesamiento alternativo a partir de un único transcrito primario, tal como sucede con el gen *tra2* de otros insectos. Cuarto, se expresa de manera constitutiva durante todo el desarrollo y la vida adulta. Y cinco, la proteína Tra2 de *Sciara* es capaz de proveer la función *Tra2*, al menos en parte, en el desarrollo sexual de *Drosophila*.

## 5.1 ANÁLISIS COMPARATIVO DE LAS PROTEÍNAS TRA2 DE INSECTOS

Las proteínas Tra2 pertenecen a la familia SR de proteínas, las cuales están involucradas en la regulación del procesamiento de los transcritos primarios, en el transporte del ARNm y en la traducción del mismo (Shepard y Hertel 2009). Se caracterizan por poseer un dominio RRM de unión al ARN y dos dominios RS de interacción con otras proteínas. Los dominios RS se caracterizan por ser ricos en dipéptidos arginina-serina.

La comparación de las proteínas Tra2, de los insectos caracterizadas hasta el momento, ha puesto de manifiesto que el dominio RRM y la región "linker" son las regiones de la proteína más conservadas, en contraste con la diversidad mostrada en los dominios RS1 y RS2, tanto por lo que respecta al número de amino ácidos que los componen, que es el mismo para el RRM (72 amino ácidos) y el "linker" (19 amino ácidos) mientras que varía para los RS, así como por el grado de similitud de los mismos (amino ácidos idénticos y cambios conservativos), que es mayor para RRM y el "linker". Esto está de acuerdo con el hecho de que el dominio RRM y el "linker" juegan un papel esencial en la función de las proteínas

Tra2. De hecho, la región "linker" está considerada como un motivo que identifica a las proteínas Tra2 frente a otras proteínas SR (McAllister *et al.*, 2000). A este respecto, es de destacar el hecho de que dos de las tres mutaciones de falta de función en el gen *tra2* de *Drosophila*, que afectan tanto la determinación sexual como la espermatogénesis, se localizan en la región del "linker" (Amrein *et al.*, 1994). Por otro lado, el dominio RRM confiere a la proteína Tra2 su capacidad de unión al ARN. Un ejemplo paradigmático de la función de Tra2 como elemento de unión al ARN es su interacción con la proteína Tra para formar el complejo Tra-Tra2 que controla el procesamiento del pre-ARNm del gen *dsx*. La proteína Tra2 es la que confiere a dicho complejo su capacidad de unión a ese pre-ARNm, ya que la proteína Tra, la cual juega el papel discriminatorio central del desarrollo sexual del cigoto, carece de dominio de unión al ARN.

Los dominios RS1 y RS2 presentan diversidad entre las proteínas Tra2, tanto por lo que respecta al número de aminoácidos que los componen como al número de dipéptidos arginina-serina (RS) que poseen. Esta variación en los dominios RS ha sido observada, también, para las proteínas Tra de los insectos caracterizadas hasta el momento (Ruiz *et al.*, 2007b). Aunque dichas regiones RS son esenciales para la función de la proteína Tra2, ya que confieren a la misma su capacidad de interaccionar con otras proteínas SR como, por ejemplo, la proteína Tra, dichas regiones parecen admitir variaciones en el número de sus dipéptidos arginina-serina, siempre y cuando la reducción de estos dipéptidos no comprometa la funcionalidad de las proteínas Tra2.

El grado de similitud entre las proteínas Tra2 no sigue las relaciones filogenéticas de las especies comparadas. Así, por ejemplo, la proteína Tra2 de *Sciara*, que es un díptero, es más similar a la de *Bombyx*, que es un lepidóptero, que a la de los otros dípteros *Ceratitis*, *Lucilia* y *Drosophila*, con la excepción de *Musca*. Dentro de los dípteros, la proteína Tra2 de *Sciara* es más similar a *Musca*, que pertenece al grupo Calyptratae, después a *Ceratitis*, que pertenece al grupo Acalyptratae, después a *Lucilia*, que pertenece al grupo Cayptratae, y después a *Drosophila*, que pertenece al grupo Acalyptratae. Esta situación contrasta con lo encontrado para el gen *dsx*, donde existe una concordancia entre el grado de similitud de este gen y las relaciones filogenéticas de las especies donde ha sido caracterizado (Ruiz *et al.*, 2007b). Esta discrepancia la explicamos por el hecho de que el gen *dsx* codifica factores de transcripción que controlan los genes de cito-diferenciación sexual, por lo que *dsx* está sujeto a una selección purificadora fuerte para mantener dicho control. Esta selección no parece ser tan fuerte en las proteínas

Tra2, pues muestran variabilidad en sus dominios RS, siendo ésta una de las características de las proteínas SR, siempre y cuando la variabilidad no comprometa su funcionalidad (Shepard y Hertel, 2009).

## 5.2 CONSERVACIÓN EVOLUTIVA DE LA FUNCIÓN DEL GEN *tra2* DE *Sciara* EN LA DETERMINACIÓN SEXUAL DE *Drosophila*

La ausencia de función *tra2* en *Drosophila* determina que los cigotos cromosómicamente hembras (XX) se desarrollen como pseudomachos, los cuales son indistinguibles de los machos normales XY, excepto que son estériles. Esto es debido a que, en ausencia de función *tra2*, el procesamiento del transcrito primario del gen *dsx* produce ARNm *dsxM*, el cual codifica la proteína DsxM que determina que el cigoto se desarrolle como macho. Aunque en esos pseudomachos se produce proteína Tra, no se forma el complejo Tra-Tra2, el cual determina que el pre-ARNm del gen *dsx* se procese de forma que de lugar a ARNm *dsxF*, el cual codifica la proteína DsxF que hace que el cigoto se desarrolle como hembra.

La expresión del gen *tra2* de *Sciara* en pseudomachos (XX) de *Drosophila* mutantes para su propio gen *tra2*, causó el procesamiento tipo hembra del pre-ARNm del gen *dsx*. Este procesamiento no se observó, sin embargo, cuando la expresión de *Tra2* de *Sciara* ocurre en los machos XY de *Drosophila* mutantes para *tra2*. Esto indica que el efecto de *tra2* de *Sciara* es dependiente de la presencia de la proteína endógena Tra de *Drosophila*, la cual está presente en los pseudomachos XX pero ausente en los machos normales XY. Por lo tanto, el gen *tra2* de *Sciara* es capaz de proveer la función *tra2* involucrada en la determinación sexual de *Drosophila*, a través del mismo mecanismo por el que lo hace la propia proteína Tra2 de esta especie; esto es, la proteína Tra2 de *Sciara* es capaz de interaccionar con la proteína Tra de *Drosophila* y determinar el procesamiento del pre-ARNm del gen *dsx* de esta especie.

Los pseudomachos que expresaban el gen *tra2* de *Sciara* producían, además del ARNm *dsxF* normal, un ARNm que tenía incorporado el intrón 3 que precede al exón 4 específico del ARNm *dsxF* de hembra. Este ARNm aberrante no produce una proteína normal DsxF. Esta retención del intrón 3 no parece ser consecuencia de un problema general del procesamiento del pre-ARNm de *dsx* en los pseudomachos que expresan el transgén *tra2* de *Sciara* porque el procesamiento

del intrón 2 (y, por lo tanto, se espera que, también, el del intrón 1) ocurre normalmente. Tampoco puede ser causado por el efecto del choque térmico a que son sometidos esos pseudomachos, pues sus hermanos XY no expresan el ARNm *dsxF* a pesar de recibir el mismo choque térmico. Finalmente, ese procesamiento aberrante no parece deberse a una expresión baja del transgén *tra2* de *Sciara* en *Drosophila*, de modo que hubiese una cantidad limitante de proteína Tra2, pues la expresión de ese transgén es lo suficiente como para inducir la letalidad de los cigotos que lo expresan.

Dado que el procesamiento del intrón 3 requiere la interacción del complejo Tra-Tra2 con sus secuencias diana localizadas en el exón específico de hembra que sigue a ese intrón, el procesamiento aberrante del pre-ARNm *dsx*, que supone la retención de dicho intrón en las moscas transgénicas de *Drosophila* que expresan el gen *tra2* de *Sciara*, se debería a una alteración en el complejo formado por la proteína Tra de *Drosophila* y la proteína transgénica Tra2 de *Sciara*. Dicho complejo tendría una eficiencia menor que el complejo formado por las propias proteínas Tra y Tra2 de *Drosophila*. En este contexto, es importante destacar tres observaciones experimentales. Primera, que la mayor divergencia entre las proteínas Tra2 de *Drosophila* y de *Sciara* se debe a las variaciones en los dominios RS, los cuales son responsables de la interacción de Tra2 con Tra (resultados de esta tesis). Segunda, que las moscas de *Drosophila* con una dosis de *tra* y una dosis de *tra2*, poseen niveles insignificantes de la isoforma aberrante de ARNm *dsxF* que retiene el intrón 3, niveles iguales a los encontrados en moscas que son portadoras de dos dosis de *tra* y dos dosis de *tra2* (Ruiz y Sánchez, 2010). Y tercera, que el mismo tipo de procesamiento aberrante del pre-ARNm *dsx* reteniendo el intrón 3 se ha encontrado en pseudomachos de *Drosophila* mutantes para su gen *tra* endógeno y que expresaban el gen *tra* del díptero *Anastrepha* (tefrítido) (Ruiz y Sánchez, 2010). Las proteínas Tra de *Anastrepha* y de *Drosophila* varían fundamentalmente en sus dominios RS (Ruiz *et al.*, 2007b).

Todos estos resultados apoyan la idea de una co-evolución molecular de las proteínas Tra y Tra2 involucradas en determinación sexual. Los cambios acumulados en dichas proteínas, después de la separación de los distintos linajes en los dípteros, harían que la interacción de estas proteínas estuviese afectada, de forma que los complejos Tra-Tra2 inter-específicos fuesen menos eficientes que los intra-específicos en su función de controlar el desarrollo sexual, a través de su control sobre el procesamiento del transcrito primario del gen *dsx*.

### 5.3 FUNCIÓN Y EVOLUCIÓN DEL GEN *tra2*

La proteína Tra2 de *Drosophila* presenta una dualidad de función. Por un lado, actúa como un activador del procesamiento del transcrito primario del gen *dsx* en las hembras y, por otro lado, actúa como un inhibidor del procesamiento de su propio transcrito primario durante la espermatogénesis en el macho adulto (ver Introducción). Mientras que la función activadora requiere de la proteína Tra, la función inhibidora es independiente de esta proteína, ya que en los machos no existe Tra. Además, las secuencias diana de la proteína Tra2 en su función inhibidora son diferentes de las secuencias diana involucradas en su función activadora.

La presencia del intrón M1 en el ARNm maduro tiene lugar solamente en la espermatogénesis del macho adulto, previniendo la formación excesiva de proteína Tra2 funcional, pues dicho intrón introduce codones de parada de la traducción, lo que da lugar a una proteína Tra2 truncada no funcional. Así pues, en la línea germinal del macho se dan los dos tipos de ARNm que codifican la proteína completa funcional y la truncada, respectivamente, representando el último tipo alrededor del 60% del total del ARNm de *tra2* (Mattox y Baker, 1991). Esta función negativa de Tra2 es dependiente de su concentración, pues las células somáticas que tienen aumentados sus niveles de proteína Tra2 funcional expresan, también, el ARNm conteniendo el intrón M1 (Qi *et al.*, 2006). El hecho de que la función inhibidora de *tra2* se de solamente en la línea germinal se debe a que el promotor que funciona en las células germinales es más activo que el que funciona en el soma (Mattox *et al.*, 1990; Amrein *et al.*, 1990; Qi *et al.*, 2006). La regulación de los niveles de proteína Tra2 es necesaria, pues niveles altos de la misma causan esterilidad del macho por alteración de la espermatogénesis (McGuffin *et al.*, 1998) y letalidad de los machos y hembras (Qi *et al.*, 2006). Por lo tanto, la función negativa de *tra2* en la línea germinal del macho adulto tiene como objetivo el reducir los niveles de proteína Tra2 funcional en este tejido.

Esta dualidad de función de Tra2 no es exclusiva de *Drosophila*, pues esta proteína actúa como activador del procesamiento del pre-ARNm de *dsx* y como inhibidor del procesamiento del pre-ARNm del gen *tra* en *Musca* (Burghardt *et al.*, 2005), *Ceratitis* (Salvemini *et al.*, 2009) y *Lucilia* (Concha y Scout, 2009). En estas especies, el gen *tra2* produce un único transcrito en el soma y en la línea



germinal de ambos sexos, no encontrándose niveles significativos del transcrito parcialmente procesado específico del testículo de *Drosophila*.

En el caso de *Sciara*, el ARNm más abundante del gen *tra2* es el que codifica para la proteína Tra2 funcional. Contrario a lo descrito para *Musca*, *Ceratitis* y *Lucilia*, en los machos pero no en las hembras de *Sciara* sí se detectan niveles de un ARNm parcialmente procesado, el transcrito *Sotra2-130*, el cual porta el exón E4a que es intrón en el transcrito *Sotra2-251* que codifica la proteína Tra2 funcional. Dicho exón contiene en *S. coprophila* y en *S. ocellaris* una y dos secuencias ISS, respectivamente, de unión de la proteína Tra2 descritas en *Drosophila*, e introduce en el ARNm *Sotra2-130* codones de parada de la traducción, lo que determina que dicho mensajero codifique para una proteína Tra2 truncada, conteniendo solamente el dominio RS1, y, por lo tanto, presumiblemente no funcional. Por lo tanto, el transcrito *Sotra2-130* de *Sciara* sería similar al transcrito *tra2-179* de *Drosophila*, el cual retiene el intrón M1. Sin embargo, mientras que el transcrito *tra2-179* es específico del testículo en *Drosophila*, el transcrito *Sotra2-130* de *Sciara* se encuentra a niveles similares en el testículo y en el soma (resto del adulto). Estos resultados podrían sugerir que la proteína Tra2 de *Sciara* tuviese una función inhibidora del procesamiento completo de su transcrito primario, como ocurre en el testículo de *Drosophila*, que redujese sus niveles y de esta forma no comprometiese ni la viabilidad ni la esterilidad de los machos (hay que recordar que la expresión incontrolada de la proteína Tra2 de *Sciara* en *Drosophila* causa letalidad de ambos sexos). Si tal función existiese, ésta estaría restringida a los machos, pues las hembras no presentan niveles significativos del ARNm *Sotra2-130*, y supondría que el promotor del gen *tra2* sería más activo en los machos que en las hembras de *Sciara*. Aunque no podemos descartar, a priori, esta posibilidad, algunas evidencias experimentales indirectas sugieren que el mayor nivel de expresión del transcrito *Sotra2-130* en los machos sería irrelevante.

Primero, la diferente activación del gen *tra2* en *Sciara* es realmente menor si se compara con el nivel de expresión del gen *tra2* en el testículo versus el resto del organismo en *Drosophila*, donde la expresión es alrededor de 14 veces mayor en el testículo (Qi *et al.*, 2006). En el caso del macho de *Sciara*, la expresión de *tra2*, que supone el transcrito *Sotra2-251* más el transcrito *Sotra2-130*, sería alrededor de 1,2 veces la expresión en la hembra, si se tiene en cuenta que la expresión del primero es similar en ambos sexos, y que la expresión del segundo en relación con el primero es de aproximadamente 0,2 en los machos.

Segundo, el gen *tra2* de *Drosophila* posee dos promotores, siendo el promotor aguas abajo el que se expresa específicamente en el testículo (Amrein *et al.*, 1990; Mattox *et al.*, 1990). En el caso de *Sciara*, como en el de *Musca* (Burghardt *et al.*, 2005), *Ceratitis* (Salvemini *et al.*, 2009) y *Lucilia* (Concha y Scout, 2009), el gen *tra2* posee un único promotor, el cual da lugar a un único transcrito primario, que es el mismo en ambos sexos, tanto en las gónadas como en el resto del organismo.

En conjunto, estos resultados sugieren que durante la evolución del linaje que dio lugar a los drosofilidos —la expresión de *tra2* es similar en *D. melanogaster* y en *D. virilis* (Chandler *et al.*, 1997)— el gen *tra2* adquirió la función inhibidora del procesamiento de su propio transcrito primario, como una solución evolutiva a la adquisición de un segundo promotor que específicamente funciona en el testículo de una manera muy activa, regulando así los niveles de proteína Tra2 funcional en esas gónadas, dado que altos niveles de esta proteína alteran la espermatogénesis. Hay que recordar que Tra2 es una proteína SR y que, por lo tanto, en exceso puede interferir con el procesamiento de los transcritos primarios y/o transporte de los mensajeros al citoplasma y/o su traducción.

El gen *tra2* de *Sciara* participa en el control del desarrollo sexual en *Drosophila*, mediante su función en el procesamiento específico de hembra del transcrito primario del gen *dsx*. ¿Está *tra2* involucrado en el desarrollo sexual de *Sciara*? En nuestro grupo, se ha clonado el gen *dsx* de *Sciara* (El gen *doublesex* de *Sciara* (Díptera, Nematocera, Sciaridae), Tesis Doctoral de Mercedes Alvarez García, Universidad Complutense de Madrid, 2009). A diferencia de otros insectos donde se ha caracterizado este gen, en *Sciara* no hay un procesamiento específico de sexo del transcrito primario de *dsx*: los transcritos que codifican para las proteínas de hembra, DsxF, y de macho, DsxM, se encuentran en ambos sexos. Además, la proteína DsxF se encuentra en cantidades similares en machos y hembras (la proteína DsxM no ha sido detectada). Esto indica que el gen *dsx* de *Sciara* no juega el papel central discriminatorio del control del desarrollo sexual que ejerce en otros insectos. Por otro lado, se han identificado secuencias similares a las que unen el complejo Tra-Tra2 en el exón 3 del pre-ARNm de *dsx*. Sin embargo, dichas secuencias no reconocen el complejo Tra-Tra2, pues el procesamiento de dicho transcrito es el mismo en hembras (que tienen proteína Tra) y en machos (que carecen de proteína Tra) de *Drosophila*. Por lo tanto, en el momento actual no podemos contestar la pregunta del requerimiento de *tra2* para el desarrollo sexual de *Sciara*.

En los organismos donde no existe una genética desarrollada y, por lo tanto, no existen mutaciones, la función de los genes que se caracterizan puede ser analizada mediante la metodología de la microinyección de RNAi. Desgraciadamente, esta metodología, aunque posible, no es factible en *Sciara*, dada la extrema fragilidad de sus huevos (L. Sánchez, comunicación personal), lo que hace técnicamente inviable microinyectar el elevado número de huevos necesarios para testar si la destrucción de la función endógena de *tra2* causa alteraciones en el desarrollo sexual de *Sciara*. Una posibilidad es la electroporación, que ha sido utilizada con éxito en *M. Domestica* (Leopold *et al.*, 1996) y en *B. mori* (Shamila y Mathavan, 1998; Guo *et al.*, 2004).

## **6. CONCLUSIONES**

1. El gen *tra2* de *Sciara ocellaris* y *Sciara coprophila* se expresa constitutivamente en hembras y machos durante todo el desarrollo y la vida adulta. Su transcrito primario sufre un procesamiento alternativo dando lugar a tres y cuatro transcritos diferentes en *S. ocellaris* y en *S. coprophila*, respectivamente. El transcrito más abundante, presente en cantidades similares en machos y hembras, tanto en las gónadas como en el resto del organismo, corresponde al que codifica la proteína Tra2 funcional. Un segundo transcrito retiene parte del intrón 3 y codifica para una proteína Tra2 truncada, presumiblemente no funcional. El resto de los transcritos, uno de *S. ocellaris* y dos más de *S. coprophila*, que codificarían, también, para una proteína Tra2 truncada, están muy poco representados. Por lo tanto, al contrario de lo que ocurre en *Drosophila*, en *Sciara* no existe un transcrito específico de testículo de macho, como ocurre en los dípteros *Musca*, *Ceratitis* y *Lucilia*.
2. La proteína Tra2 de *Sciara* posee las características que definen a la familia de proteínas SR; esto es, contiene un dominio RRM de unión a ARN, flanqueado por dos dominios RS1 y RS2 ricos en dipéptidos de arginina-serina, y una región "linker" que une el dominio RRM con el RS2.
3. La comparación de la proteína Tra2 de *Sciara* con la de otros insectos evidencia que el mayor grado de conservación ocurre en el dominio RRM y en la región "linker", mientras que los dominios RS1 y RS2 presentan variabilidad en cuanto al número de aminoácidos que los componen como al número de dipéptidos arginina-serina que poseen.
4. La proteína Tra2 de *Sciara* es capaz de inducir el procesamiento específico de hembra del transcrito primario del gen *dsx* de *D. melanogaster*, mimetizando la función de la propia proteína Tra2 de *Drosophila*. Esto indica que la proteína Tra2 de *Sciara* es capaz de formar un complejo con la proteína Tra de *D. melanogaster*. Sin embargo, este complejo no parece ser tan eficiente como el propio complejo Tra-Tra2 de *D. melanogaster* en inducir ese procesamiento, lo que sugiere la existencia de co-evolución entre las proteínas Tra y Tra2.

## **7. BIBLIOGRAFÍA**

ÁLVAREZ, M. (2009). El gen *doublesex* de *Sciara* (Díptera, Nematocera, Sciaridae). Tesis doctoral. Universidad Complutense de Madrid, Madrid.

AMREIN, H., GORMAN, M., NOTHIGER, R. (1988). The sex-determining gene *tra2* of *Drosophila* encodes a putative RNA binding protein. *Cell*. 55(6): 1025-1035.

AMREIN, H., HEDLEY, M. L., MANIATIS, T. (1994) The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by *transformer 2*. *Cell*. 76(4):735-746.

AMREIN, H., MANIATIS, T., NÖTHINGER, R. (1990). Alternative spliced transcripts of the sex determining gene *tra2* of *Drosophila* encode functional proteins of different size. *EMBO J.* 9: 3619-3629.

BACHILLER, D. and SÁNCHEZ, L. (1991). Production of XO clones in XX females of *Drosophila*. *Genet Res.* 57: 23-28

BELL, L. R., MAINE, E. M., SCHEDL, P. and CLINE, T. W. (1988). *Sex-lethal*, a *Drosophila* gene determination switch gene, exhibits sex-specific RNA splicing and sequence similar to RNA binding proteins. *Cell* 55: 1037-1046.

BELOTE, J. M. and BAKER, B. S. (1993). The dual function of a sex determining gene in *Drosophila melanogaster*. *Dev. Biol.* 95: 512-517.

BELOTE, J. M., MCKEOWN, M., BOGGS, R. T., OHKAWA, R. and SOSNOWSKI, B. A. (1989). Molecular genetics of *transformer*, a genetic switch controlling sexual differentiation in *Drosophila*. *Dev Genet.* 10(3): 143-154.

BEYE, M., HASSELMAN, M., KIM FONDRAK, M., PAGE, R. E. Jr. and OMHOLT, S. S. (2003). The gene *cds* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* 114: 419-429.

BILLETER JC, VILLELLA A, ALLENDORFER JB, DORNAN AJ, RICHARDSON M, GAILEY DA, GOODWIN SF. (2006). Isoform-specific control of male neuronal differentiation and behavior in *Drosophila* by the *fruitless* gene. *Curr Biol.* 16:1063-1076.

- BILLETER, J.C., GOODWIN, S.F., O'DELL, K.M. (2002). Genes mediating sex-specific behaviors in *Drosophila*. *Adv Genet.* 47: 87-116.
- BLACK DL. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem.* 72:291-336.
- BLENCOWE BJ. (2000). Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem Sci.* 25(3):106-110
- BOGGS, R.T., GREGOR, P., IDRIS, S., BELOTE, J.M. and MCKEOWN, M. (1987). Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the *transformer* gene. *Cell* 50(5): 739-747.
- BOOP, D., CALHOUN, G., HORABIN, J.I., SAMUELS, M., SCHEDL, P. (1996). Sex-specific control of *Sex-lethal* is a conserved mechanism for sex determination in the genus *Drosophila*. *Development* 122(3): 971-982.
- BRAND, A.H., and PERRIMON, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118(2):401-415.
- BULL, J.B. (1983). Evolution of sex determining mechanisms. *The Benjamin/Cummings Publishing Company, Inc, Menlo Park, California, USA.*
- BURGHARDT, G., HEDIGER, M., SIEGENTHALER, C., MOSER, M., DÜBENDORFER, A., BOOP, D. (2005). The *transformer2* gene in *Musca domestica* is required for selecting and maintaining the female pathway of development. *Dev genes Evol.* 215(4): 165-176.
- CHANDLER DS, QI J, MATTOX W. (2003). Direct repression of splicing by *transformer-2*. *Mol Cell Biol.* 23(15):5174-5185.
- CHANDLER, D., MCGUFFIN, M.E., MATTOX, W.(2001).Functionally antagonistic sequences are required for normal autoregulation of *Drosophila tra2* pre-mRNA splicing. *Nucleic Acids Res.* 29(14):3012-3019.



- CHANDLER, D., MCGUFFIN, M.E., PISKUR, J., YAO, J., BAKER, B.S., MATTOX, W. (1997). Evolutionary conservation of regulatory strategies for the sex determination factor *transformer-2*. *Mol Cell Biol.* 17 (5): 2908-2919.
- CHASE, B.A. and BAKER, B.S. (1995). A genetic analysis of *intersex*, a gene regulating sexual differentiation in *Drosophila melanogaster* females. *Genetics* 139. 1649-1661.
- CHEN SL, DAI SM, LU KH, CHANG C. (2008). Female-specific *doublesex* dsRNA interrupts yolk protein gene expression and reproductive ability in oriental fruit fly, *Bactrocera dorsalis* (Hendel). *Insect Biochem Mol Biol.* 38(2):155-65.
- CHO S, HUANG ZY, ZHANG J. (2007). Sex-specific splicing of the honeybee *doublesex* gene reveals 300 million years of evolution at the bottom of the insect sex-determination pathway. *Genetics.* 177(3):1733-41
- CLINE, T. W. (1993). The *Drosophila* sex determination signal: How does flies count to two?. *Trends Genet* 9(11): 385-390.
- CLINE, T.W. (1984). Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state. *Genetics* 107: 231-277.
- CONCHA, C., SCOTT, M.J., (2009). Sexual development in *Lucilia cuprina* (Diptera, Calliophoridae) is controlled by the *transformer* gene. *Genetics* 182:785-798.
- DE SAINT PHALLE B, SULLIVAN W. (1996). Incomplete sister chromatid separation is the mechanism of programmed chromosome elimination during early *Sciara coprophila* embryogenesis. *Development.* 122(12):3775-3784.
- DÜBENDORFER A, HEDIGER M, BURGHARDT G, BOPP D. (2002). *Musca domestica*, a window on the evolution of sex-determining mechanisms in insects. *Int J Dev Biol.* 46(1):75-79.
- DUBOIS, A.M. (1933). Chromosome behavior during cleavage in the eggs of *Sciara coprophila* (diptera) in relation to the problem of sex determination. *Z. Zellforsch. Mikrosk. Anat.* 19: 595-614.

- DUFFY, J. B. (2002). Gal4 System in *Drosophila*: A Fly Geneticist's Swiss Army Knife. *Genesis* 34: 1-15
- ERICKSON JW, QUINTERO JJ. (2007). Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in *Drosophila*. *PLOS BIOL.* 5(12)
- GAILEY, D.A., BILLETER, J.C., LIU, J.H., BAUZON, F., ALLENDORFER, J.B., GOODWIN, S.F. (2006). Functional conservation of the *fruitless* male sex-determination gene across 250 Myr of insect evolution. *Mol Biol Evol.* 23(3): 633-643.
- GARRET-ENGELE, C.M., SIEGAL, M.L., MANOLI, D.S., WILLIAMS, B.C., LI, H. and BAKER, B.S. (2002). *intersex*, a gene required for female sexual development in *Drosophila*, is expressed in both sexes and function together with *doublesex* to regulate terminal differentiation. *Development* 129: 4661- 4675.
- GERBI, S.A. (1986). Unusual chromosome movements in sciarid flies. *Results Probl Cell Differ.* 13: 71-104.
- GODAY C, ESTEBAN MR. (2001). Chromosome elimination in sciarid flies. *Bioessays.* 23(3):242-250
- GOODWIN, S.F., TAYLOR, B.J., VILLELA, A., FOSS, M., RYNER, L.C., BAKER, B.S. and HALL, J.C. (2000). Molecular defects in the expression of the *fruitless* gene of *Drosophila melanogaster* caused by aberrant splicing in P-element insertional mutants. *Genetics* 154: 725-745.
- GORALSKI, T.J, EDSTRÖM, J.E, BAKER, B.S. (1989). The sex determination locus *transformer-2* of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. *Cell* 56(6):1011-1018.
- GUO X.Y, DONG L, WANG S.P, GUO T.Q, WANG J.Y, LU C.D. (2004). Introduction of foreign genes into silkworm eggs by electroporation and its application in transgenic vector test. *Acta Biochim Biophys Sin.* 36(5):323-30.

GRAVELEY BR, HERTEL KJ, MANIATIS T. (2001). The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA*. 7(6):806-818.

GRAVELEY BR. (2000). Sorting out the complexity of SR protein functions. *RNA*. 6(9):1197-1211.

HEDIGER, M., BURGHARDT, G., SIEGENTHALER, C., BUSER, N., HILFIKER-KLEINER, D., DÜBENDORFER, A., BOOP, D. (2004). Sex determination in *Drosophila melanogaster* and *Musca domestica* converges at the level of the terminal regulator *doublesex*. *Dev Genes Evol*. 214(1): 29-42.

HEDIGER M, HENGGELE C, MEIER N, PEREZ R, SACCONI G, BOPP D. (2010). Molecular characterization of the key switch f provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly. *Genetics*. 184(1):155-170.

HEINRICHS, V., RYNER, L.C., BAKER, B.S. (1998). Regulation of sex-specific selection of *fruitless* 5' splice sites by *transformer* and *transformer-2*. *Mol. Cell. Biol*. 18: 450-458.

HOSHIJIMA, K., INOUE, K., HIGUCHI, I., SAKAMOTO, H. and SHIMURA, Y. (1991). Control of *doublesex* alternative splicing by *transformer* and *transformer-2* in *Drosophila*. *Science* 252(5007): 833-836.

KUHN, S., SIEVERT, V., TRAUT, W. (2000). The sex-determining gene *doublesex* in the fly *Megaselia scalaris*: conserved structure and sex-specific splicing. *Genome* 43(6): 1011-1020.

KULATHINAL, R.J., SKWAREK, L., MORTON, R.A., SINGH, R.S. (2003). Rapid evolution of sex-determining gene: *transformer*: structural diversity and rate heterogeneity among sibling species of *Drosophila*. *Mol Biol Evol*. 20(3): 441-452.

LAGOS, D., KOUKIDOU, M., SAVAKIS, C., KOMITOPOULOU, K. (2007). The transformer gene in *Bactrocera oleae*: the genetic switch that determines its sex fate. *Insect Mol Biol*. 16(2): 221-230.

LAGOS, D., RUIZ. M.F., SÁNCHEZ, L. KOMITOPOULOU, K. (2005). Isolation and characterization of the *Bactrocera oleae* genes orthologous to the sex determining *Sex-lethal* and *doublesex* genes of *Drosophila melanogaster*. *Gene* 348: 111-121.

LEOPOLD R.A, HUGHES K.J, DEVAULT J.D. (1996). Using electroporation and a slot cuvette to deliver plasmid DNA to insect embryos. *Genet Anal.* 12(5-6):197-200.

LI, H. and BAKER, B.S. (1998). *hermaphrodite* and *doublesex* function both dependently and independently to control various aspects of sexual differentiation in *Drosophila*. *Development* 125: 2641-2651.

LINDSLEY, D.L. and ZIMM, G. (1992). The genome of *Drosophila melanogaster*. San Diego, California. Academia Press.

MANIATIS, T. FRITSCH, E.F. and SAMBROOK, J. (1982). Molecular cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*.

MATTOX, W., AND BAKER, B.S. (1991). Autoregulation of the splicing of transcripts from the *transformer-2* gene of *Drosophila*. *Genes Dev.* 5:786-796.

MATTOX, W., MCGUFFIN, M.E., BAKER, B.S. (1996). A negative feedback mechanism revealed by functional analysis of the alternative isoforms of the *Drosophila* splicing regulator *transformer-2*. *Genetics* 143:303-314.

MATTOX, W., PALMER, M. J. AND BAKER, B. S. (1990). Alternative splicing of the sex determination gene *transformer-2* is sex-specific in the germ line but not in the soma. *Genes Dev.* 5:789-805.

MCALLISTER B.F, MCVEAN G.A. (2000). Neutral evolution of the sex-determining gene *transformer* in *Drosophila*. *Genetics*. 154(4):1711-1720.

MCGUFFIN, M.E., CHANDLER, D., SOMAIYA, D., DAUWALDER, B., MATTOX, W. (1998). Autoregulation of *transformer-2* alternative splicing is necessary for normal male fertility in *Drosophila*. *Genetics*. 149(3):1477-1486.

- MEISE, M., HILFIKER-KLEINER, D., DÜBENDOFER, A., BRUNNER, C., NÓTHINGER, R., BOOP, D. (1998). *Sex-lethal*, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*. *Development* 125(8): 1487-1494.
- METZ, C.W. (1938). Chromosome behavior, inheritance and sex determination in *Sciara*. *Amer. Nat.* 72: 485-520.
- MORI, L., DESSEN, E.M., PERONDINI, A.L. (1979). A gene that modifies the sex ratio in a bisexual strain of *Sciara ocellaris*. *Heredity* 42(3):353-357.
- MÜLLER-HOLTKAMP, F. (1995). The *Sex-lethal* gene homologue in *Chrysomya rufifacies* is highly conserved in sequence and exón-intron organization. *J Mol Evol.* 41(4): 467-477.
- NIGRO, R.G., CAMPOS, C.C., PERONDINI, A.L.P. (2007). Temperatura and the progeny sex-ratio in *Sciara ocellaris* (Diptera, Sciaridae). *Genetics and Molecular Biology* 30: 152-158.
- NIIMI, T., SAHARA, K., OSHIMA, H., YASUKOCHI, Y., IKEO, K., TRAUT, W. (2006). Moleculalr cloning and chromosomal localization of the *Bombyx Sex-lethal* gene. *Genome* 49(3): 263-268.
- NIU BL, MENG ZQ, TAO YZ, LU SL, WENG HB, HE LH, SHEN WF. (2005). Cloning and alternative splicing analysis of *Bombyx mori transformer-2* gene using silkworm EST database. *Acta Biochim Biophys.* 37(11):728-36.
- OHBAYASHI, F., SUZUKI, M.G., MITA, K., OKANO, K., SHIMADA, T. (2001). A homologue of the *Drosophila doublesex* gene is transcribed into sex-specific mRNA isoforms in the silkworm, *Bombyx mori*. *Comp Biochem Physiol B Bioche, Mol Biol.* 128(1): 145-158.
- O'NEIL, M.T., BELOTE, J.M. (1992). Interspecific comparision of the *transformer* gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. *Genetics* 131(1): 113-128.

- PANE, A., DE SIMONE, A., SACCONI, G., POLITO, C. (2005). Evolutionary conservation of *Ceratitis capitata transformer* gene function. *Genetics* 171(2): 615-624.
- PANE, A., SALVEMINI, M., M., DELLI BOVI, P., POLITO, C., SACCONI, G. (2002). The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development* 129(15): 3715-3725.
- PENALVA, L.O., SAKAMOTO, H., NAVARRO-SABATÉ, A., SAKASHITA, E., GRANADINO, B., SEGARRA, C., SANCHEZ, L. (1996). Regulation of the gene *Sex-lethal*: a comparative analysis of *Drosophila melanogaster* and *Drosophila subobscura*. *Genetics* 144(4): 1653-1664.
- PERONDINI, A.P.L. and DESSEN, E.M.B. (1985). Polytene chromosomes and the puffing patterns in the salivary glands of *Sciara ocellaris*. *Rev. Brasil. Genet.* 8: 465-478.
- PULTZ, M. and BAKER, B. (1995). The dual role of *hermaphrodite* in the *Drosophila* sex determination regulatory hierarchy. *Development* 121: 99-111.
- QI, J., SU, S., MATTOX, W. (2007) .The *doublesex* splicing enhancer components Tra2 and Rbp1 also repress splicing through an intronic silencer. *Mol Cell Biol.* 27(2):699-708.
- QI J, SU S, MCGUFFIN M.E, MATTOX W. (2006). Concentration dependent selection of targets by an SR splicing regulator results in tissue-specific RNA processing. *Nucleic Acids Res.* 34(21):6256-6263.
- RAMOS-ONSINS, S., SEGARRA, C., ROZAS, J., AGUADÉ, M. (1998). Molecular and chromosomal phylogeny in the obscura group of *Drosophila* inferred from sequences of the *rp49* gene region. *Mol Phylogenet Evol.* 9(1):33-41.
- RIDEOUT, E.J., BILLETER, J.C. and GOODWIN, S.F. (2007). The sex-determination gene *fruitless* and *doublesex* specify a neural substrate required for courtship song. *Current Biology* 17: 1473-1478.

- RUIZ, M. F. and SÁNCHEZ, L. (2010). Effect of the gene *transformer* of *Anastrepha* on the somatic sexual development of *Drosophila*. *Int. J. Dev. Biol.* 54: 627-633.
- RUIZ, M.F., GODAY, C., GONZALEZ, P., SANCHEZ, L. (2003). Molecular analysis and developmental expression of the *Sex-lethal* gene of *Sciara ocellaris* (Diptera Order, Nematocera suborder). *Gene Exp. Patterns* 3(3): 341-346.
- RUIZ, M.F., MILANO, A., SALVEMINI, M., EIRIN-LOPEZ, J.M., PERONDINI, A.L., SELIVON, D., POLITO, C., SACCONI, G., SANCHEZ, L. (2007,a) The gene *transformer* of *Anastrepha* fruit flies (Diptera, tephritidae) and its evolution in insects. *PLoS One*. 2(11).
- RUIZ, M.F., EIRIN-LOPEZ, J.M., STEFANI, R.N., PERONDINI, A.L., SELIVON, D. SANCHEZ, L. (2007,b). The gene *doublesex* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. *Dev Genes Evol.* 217(10):725-731.
- RYNER, L.C., GOODWIN, S.F., CASTRILLON, D.H., ANALD, A., VILLELA, A., BAKER, B.S., HALL, J.C., TAYLOR, B.J. and WASSERMAN, S.A. (1996). Control of male sexual orientation in *Drosophila* by the *fruitless* gene. *Cell* 87: 1079-1089.
- SACCONI G, PANE A, POLITO LC. (2002) Sex determination in flies, fruitflies and butterflies. *Genetica*. 116(1):15-23.
- SACCONI, G., PELUSO, I., ARTIACO, D., GIORDANO, E., BOOP, D., POLITO, L.C. (1998). The *Ceratitis capitata* homologue of the *Drosophila* sex determining gene *Sex-lethal* is structurally conserved, but not sex-specifically regulated. *Development* 125(8): 1495-1500.
- SALVEMINI, M., ROBERTSON, M., ARONSON, B., ATKINSON, P., POLITO, L.C., SACCONI, G. (2008). *Ceratitis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination. *Int J Dev Biol.* 53(1):109-120.
- SALZ, H.K., MAINE, E.M., KEYES, L.N., SAMUELS, M.E., CLINE, T.W. and SCHEDL, P. (1989). The *Drosophila* femalespecific sex-determination gene, *Sex-lethal*, has

stage-, tissue-, and sex-specific RNAs suggesting multiple models of regulation. *Genes Devel.* 3: 708-719.

SAMBROOK, J., FRISTSCH, E.F. and MANIATIS, T. (1989). Molecular cloning: a Laboratory Manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.*

SHAMILA Y, MATHAVAN S. (1998). Sperm-mediated gene transfer in the silkworm *Bombyx mori*. *Arch Insect Biochem Physiol.* 37(2):168-177.

SÁNCHEZ, L. (2008). Sex determining mechanism in insects. *Int J Dev Biol.* 52(7): 837-856. Review.

SÁNCHEZ, L., GORFINKIEL, N. and GUERRERO, I. (2005). Sex determination and the development of the genital disc. In *Comprehensive Molecular Insect Science*, Vol.1, GILBERT, L.I., IATROU, K and GILL, S.S. (eds.) Elsevier Pergamon Oxford, UK, pp. 1-38.

SANCHEZ L, PERONDINI ALP. Sex determination in sciarid flies: a model for the control of differential X-chromosome elimination. *J. Theor Biol.* 197(2):247-259.

SÁNCHEZ, L. and NÖTHINGER, R. (1983). Sex determination and dosage compensation in *Drosophila melanogaster*: production of male clones in XX females. *EMBO J.* 2: 485-491.

SCALI, C., CATTERUCCIA, F., LI, Q., CRISANTI, A. (2005). Identification of sex-specific transcripts of the *Anopheles gambiae doublesex* gene. *J Exp Biol.* 208(Pt 19): 3701-3709.

SERNA, E., GORAB, E., RUIZ, M.F., GODOY, C., EIRIN-LOPEZ, J.M., SANCHEZ, L. (2004) The gene *Sex-lethal* of the Sciaridae family (order Diptera, suborder Nematocera) and its phylogeny in dipteran insects. *Genetics.* 168(2):907-921.

SHARP, P.A., SUGDEN, B., SAMBROOK, J. (1973). Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose ethidium bromide electrophoresis. *Biochemistry.* 12(16): 3055-3065.



- SHEARMAN, D.C., FROMMER, M. (1998). The *Bactrocera tryoni* homologue of the *Drosophila melanogaster* sex-determination gene *doublesex*. *Insect Mol Biol.* 7(4): 355-366.
- SHEPARD, P.J., HERTEL, K.J. (2009). The SR protein family. *Genome Biol.* 10(10):242.
- SHIRANGI, T.R., TAYLOR, B.J. and MCKEOWN, M. (2006). A double- switch system regulates male courtship behavior in male and female *Drosophila melanogaster*. *Nature Genetics* 38: 1435-1439.
- SIEGAL, M.L., BAKER, B.S. (2005). Functional conservation and divergence of *intersex*, a gene required for female differentiation in *Drosophila melanogaster*. . *Dev Genes Evol.* 215(1): 1-12.
- SIEVERT, V., KUHN, S., PAULULAT, A., TRAUT, W. (2000). Sequence conservation and expression of the *sex-lethal* homologue in the fly *Megaselia scalaris*. *Genome* 43(2): 382-390.
- SIEVERT, V., KUHN, S., TRAUT, W. (1997). Expression of the sex determining cascade genes *Sex-lethal* and *doublesex* in the phorid fly *Megaselia scalaris*. *Genome* 40(2): 211-214.
- SUZUKI, M.G., OHBAYASHI, F., MITA, K., SHIMADA, T. (2001). The mechanism of sex-specific splicing at the *doublesex* gene is different between *Drosophila melanogaster* and *Bombyx mori*. *Insect Biochem Mol Biol.* 31(12): 1201-1211.
- TIAN, M., MANIATIS, T.(1992). Positive control of pre-mRNA splicing in vitro. *Science* 256(5054):237-240.
- TIAN, M. and MANIATIS, T. (1993). A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. *Cell* 74: 105-114.

VALCÁRCEL, J., SINGH, R., ZAMORE, P.D. and GREEN, M.R. (1993). The protein *Sex-lethal* antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. *Nature* 362(6416): 171-175

WATERBURY, J.A., JACKSON, L.L. and SCHEDL, P. (1999). Analysis of the Doublesex female protein in *Drosophila melanogaster*. Role in sexual differentiation and behavior and dependence of Intersex. *Genetics* 152: 1653-1667.

WHITE, M.D.J. (1973). Animal cytology and evolution. 3<sup>nd</sup> end. *Cambridge University Press, Cambridge, UK*.

WU, J.Y. and MANIATIS, T.(1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*. 75(6): 1061-1070.

## **8. ANEXOS**

5' GTGAAGCTGT GTCTATTGT ACAATTGCAA TTATAAAAGC GAGCCAAGAA CCGCTATAAA TACATCTCCA  
 GTGTCATTAA TCGAATACAG AATAATGTCT CGAAGGGTAA GCCAGCTCGT TTTGTCTTCT TTTTTTATTT  
 CATAAACACA CCAATTCTTC AAGATATGAT AAATTCTCAA AATTAAATGC AATCGAAGAC ATCGGTTGTT  
 TTAGAAAAAC CATAGCAACA CACGGCTTTT GTTCATGGGC CTTTGTTTAA ATTTGTTTTT TATTCATTGA  
 TCAAACGAAA GCCACGCTTC ACAGCGATAA CAAGCCGCAA CCGGAGTTGG TGTTTAATCT TCAAATAAAT  
 TAAATCTTTC CGATTATTAG AAATGAGTTT GCTTGTAAAT CAATTTTTGT TCTCCCTTAA CCGTATAAGT  
 TGTGCAAAT CGATTAAAA ACGAAACTAT TGTTTTAGAT GTCACGTTTC AGAAGTCGGA GTTACGAAGT  
 TCACAGACGA CCTATAGAAC CTCCTTCGCC TCCATCCGTC GGCAGGAAGT ACAGAGATGA GTAAGTGACT  
 TGAATTTTTTA TTTGGGTTGG GTGCATTTGT ATAGTGATTC GATTTCATT GGGAGTCACA TAAAGTTCTT  
 TTCAGAAAGC TGATAAATGA ATATTCCATT AAGTTGTGCG TATGAACGTC TCATTGGCAA TGTGTATCTC  
 AGCAAGCACT TTTTCTTGTT GTCGAACTTT TTTTCTTCT AAAAATGCCC TTCTCTAATC ATTAAGATAC  
 AGACACAAT TTCGTTACGT ATTTCCCGGT TTACGTACAA TAGAGATTCT ATCACTATGA TCGACGCTCG  
 TCATCAATAA GCTCAAAGTA ACCGCTATGC ACAGTGACT TTGTTTGTAC CGACGAAATA AGAATTGACA  
 CTGTGTTAAA GCTCTGTACA ACTTAGTTTC TCATGAACCT GACCTGCGCT AGATCCAGAC TACTGTTCTC  
 CGAATTTTCA ATTTCCGGCTC TCTGAATTTT CAATGCGTGT CACCAAACGT GAAAAGAGCT CCTACACGGG  
 GTATTTCACT ATGTGGTTTA TCGCGTTAAA CTCTTTGCAA AGAGACCACA TGATTGCTGT GTCGGGCTGT  
 TACACTGTTT AGTGAGTGT TATCGCCTGT GTACACGCTG TCACTCTACT TTTGACACTG AAGGTCTCTA  
 TCTCGAACGT AGCTAGTTTA TCAACGAACC GGTCTGATT GGAAGTGAAG GCTTTCGTGC TAGAGTAATC  
 AAACTCGAAA GGTGGAGTGA TGGTACCTGC TGAGTGGTGT TTCTGCCTGG CTCAGATTTC ATGTTGGCCA  
 AATGCAGAGA GAGTTTACGT TCGTATGGGC TAGTTCTGTT AAATCTCGGT ATTTCTCGGT GTGTTTGGGA  
 TAATAACTTG CTTATCTCTT ATGTCGAGCG ACATATGAAT TACAAGCTTT GTCAATCTAA CTGGCAAAGA  
 TTGTGATCAG GGCAAAATG TTGTTCTTTC GGCTTTTATT GCGTAAGGC CTGAGTGGTC AGTTGACATT  
 ATTATCATGT CAGTTACAAT CTAATGAGGC GGGATTGACA CCAGCCACTC TCACAAGATC TGTGCCCATT  
 GATCAGAAAT GATCTGCCCT AATTGATGAA ATCTACCTGA CCTGAAAACA AAACCAACCT GAAACGATCA  
 ATATTGACCG GAGATGAATG TAACCAAGGT CAAGTCAGGC CGAGTTTATT TCCAAAAGTG ACATCTCAAG  
 CCCAATCAAC GATCTTAATT CAAATCTGTT ATTTCTTCCA GCGACAGCCG CAGTGGTTCT CGAGAAGTGA  
 CCCTTCGAAA GAAGCGAAGA ATGTCATATG ATCGGTCAAG ATCAGCATCT CGTTCACGAT CAATGGGACG  
 ATACGCTCAC AGATCAAGCA GATATGTGCG AAGTCGATCT CGGTAAGTTA TGATTACGAA AACAGAAGAT  
 TTGTTTTTTT CGGTTGCAAG CAACGTGTAA TCAATTGAAA ATAAGAAAAT TTTCATTCTC AACTGCCGAA  
 GTGTAGGATG TAATTTTTTT TCTTTGTCTT TTTTTTCTC TGCACTGAAG GAAGAAATAT CCAAAGAGA  
 AAAAATATTT TCAAAGCACG GGAAACACTT CAATCAATTT AATAAAAAAA AGTCAACCCG AAATCTATAA  
 TTCAACTTTA ACCAAAACCA ATTCTACAAC AAAAACCAGC TCAATTACAC GAACATAGTC CCCTCAAGAA  
 ATATCATTTT TTGTGATTTG AAATTCAGCA ACCTAACCCA AGGCCTGATG ATAAAACAAT ACAAACAAAC  
 GATCGAAGAA TGTGTTTGCG CTCTTGTTCT TTTGATGAAG GATCAAACCC AACGAATTAT AACTCGAAG  
 CAACCACAAG CATAACAAA AATTACCGAA ACGAGCACAC GCATTACAAA CTGACATCGC CGCCGTCCTT  
 CGATCATAAT TTGTGAAAAC GTGTCAAACC ATCGAATAAA TCCAAACAAA AATCTTTTCT TCCACAAAAA  
 AAGGTCGTAT TCGCCGTATA AGCGTCGCAG CGATCGTCGC AGATCCTACT CCAGAAGCCC GTACAATGAT  
 CGCCACAGAC ATGTGGGCAC ACGTGTAAGT TATCACAAC AATTTTTGGT AGTGTGCGT CGAAAGTCAT  
 TGTGTCTTTT GATTTTCTTT TAAAGGAAAA GCCGAGAAAG AGCCGTTGCC TTGGTGTGTT TGGTCTAAGT  
 GTCTACACCA CGAAAGATCA AATCGTGCAG ATATTCTCCA AGTATGGTCC AATTGAACAT GTTCAAGTGG  
 TTGTCGATGC GCGGTGAGT AGTCGTAACA TTTTGTTCCT TTACAATTGA TGCCACAGCT CCGCAGAAGT  
 AGTTTGAAAT TTTAAACTA CTTCTCGTCA AGTGGAACGT GGCTAGATTA AAGACCATT GAGAGAAATA  
 GAACGGGACT TCCTGTAGCC TTGGAGTCTC ATAAAGGCCT CACAAAACAA AACGTAATTT TTTGAACGAT  
 TCTGTTAGTC GGTGGGATG ATAGACTTTT CACGTTACAC CATATTGTTT TCCACATACC TCGTCTTTTT  
 CTGAAGATTC TTCAGGATTT CTCGATCCTC GATGAGTTTG CACGAAACT TTGCCCTAAT TTCGCTTTCT

```

ATTCGTCAAT CAGCGAATTT GTAGTCAATT TGACTTTTCC TCCGACCAGA CGACGAGCTT TGCCACAATC
GCCTGATAAA TCTTATGGTG AACTTATCAT CACAAGACCC ATGAAACGAA CGACCGTTGT GATTCTTCTT
CACTAGTCTG AGTCTAGGAC TCACCGATTT GTTGGATCGT TCATAAAACC CCATCTCCAA CTGCGCTCAG
ACCATACTCA ACGATTGTAA TCTTAAACAA ACATTTCGATT CCAGACAGGG CGTTCGCGTG GCTTTTGTTT
CGTCTACTTT GATCGCACCG ATGACGCTAC CGTTGCCAGA GACCATTGCA CCGGCATGGA AATCGATAAC
CGTCGCATTC GTGTAGATTA TTCGATAACA TTGCGTCCGC ACGAACCAAC ACCGGGCATT TACAAAGGAC
GCTCGTCACG GTCGCACCAC GAAAGAGATC GAAATGACGG TCGTCGACGA ACACCACCAT CACCGTACCG
AGAACGTGCG CGTCGTTACA GTAGATCACG TTCGTATTCA CGTGAACGGA AGTAATTTTT CGCGGGCACT
TTTCTTTTCT CCTTCTTTT TTCTTCTACG ATAATTTTGT AAGATTTAGT TTGTTAAAAA GCACAAAAGA
TAGCGGGACC ACAAATTTGT GGTGTTTGT TCAATTGGTT TACCGAAAAT TAAAAAGAAA AAAGAGAATT
CGGTATGCCG AAGGTGAGAC TAATCCGGGA CATTTGTGGA ATTGTGTTGA ATCTTTAATT GTTTTTAAGA
ATTTTCGCCA TGTCTTTTGA CCATTTTCCT TTATTACTAT TGATAGAAGA CATTTTACTA TGAAGAATTA
TAGACTTGCC AGTTAACTGC TTACCATGAA GCAGATGTAT GAAATTTTTC TTGGTTTTAT TTTACAAAAG
CAAATTTCCG TAAATCGGAA ATCAAAGTCG CCTTAAACAA TTCGAAACAC CAAAAGCATG GCAAATTAGA
TGACACACAG ACATTTAAAT TAATTGAAAT TAAAACAAAC TGAGGACACC ATTCAAAATG GTGGATACTT
TTAGGCTTAA GACGCGGTGT TTTTCCTTCT TTTCTTTTTT TTGGTTTGGA GCGTTCAACT GAAACTATAA
ATTGTGATTA AAAATTGATT TGCATTAATG AAATTTTGTT GGATAAGGCG AGGAACATTC AATTAATTAA
TTAAAAGAAA ATAAATTAAT TTCGTTTGGA TATACACTCT CACACTACCT AACAAATTGT TTATTTTTTT
TTGGAACGA TTTTCGTTTG TAATTAGCAC ATTAAAAAGA AACTTAGTT GTAATTTGAT TTAACAAACA
AATATTGAAA TGAAACAACC CACGATCGAT TCGATCAAAA AAATAATTTA GAATTAAGAC AACACTAAAA
TATTATAGAG AGAGATGATT GCAATAAAAT GTGCGAGTAT TTTTGAAAAT C 3'

```

**Secuencia de nucleótidos correspondiente al gen *tra2* de *S. ocellaris*.** Los exones aparecen en color verde y los intrones en color rojo. La secuencia subrayada corresponde al intrón que se retiene parcialmente en alguna de las formas (E4a). Los recuadros negros señalan las secuencias ISS. En azul se indican las regiones 5' y 3' UTR. La región 3' UTR que aparece coloreado en un tono más oscuro corresponde a la del transcrito más largo.

```

5' G TGA AGC TGT GTC CTA TTG TAC AAT TGC AAT TAT AAA
AGC GAG CCA AGA ACC GCT ATA AAT ACA TCT CCA GTG TCA TTA ATC GAA TAC AGA ATA
M S R R M S R S R S Y E V H R R P
ATG TCT CGA AGG ATG TCA CGT TCG AGA AGT CGG AGT TAC GAA GTT CAC AGA CGA CCT
I E P P S P P S V G R K Y R D D D S R
ATA GAA CCT CCT TCG CCT CCA TCC GTC GGC AGG AAG TAC AGA GAT GAC GAC AGC CGC
S G S R E V T L R K K R R M S Y D R S
AGT GGT TCT CGA GAA GTG ACC CTT CGA AAG AAG CGA AGA ATG TCA TAT GAT CGG TCA
R S R S R S R S M G R Y A H R S S R Y
AGA TCA CGA TCT CGT TCA CGA TCA ATG GGA CGA TAC GCT CAC AGA TCA AGC AGA TAT
V R S R S R S Y S P Y K R R S D R R R
GTG CGA AGT CGA TCT CGG TCG TAT TCG CCG TAT AAG CGT CGC AGC GAT CGT CGC AGA
S Y S R S P Y N D R H R H V G T R R K
TCC TAC TCC AGA AGC CCG TAC AAT GAT CGC CAC AGA CAT GTG GGC ACA CGT GAA AAG
P R K S R C L G V F G L S V Y T T K D
CCG AGA AAG AGC CGT TGC CTT GGT GTG TTT GGT CTA AGT GTC TAC ACC ACG AAA GAT
Q I V Q I F S K Y G P I E H V Q V V V
CAA ATC GTG CAG ATA TTC TCC AAG TAT GGT CCA ATT GAA CAT GTT CAA GTG GTT GTC
D A A T G R S R G F C F V Y F D R T D
GAT GCG GCG ACA GGG CGT TCG CGT GGC TTT TGT TTC GTC TAC TTT GAT CGC ACC GAT
D A T V A R D H C T G M E I D N R R I
GAC GCT ACC GTT GCC AGA GAC CAT TGC ACC GGC ATG GAA ATC GAT AAC CGT CGC ATT
R V D Y S I T L R P H E P T P G I Y K
CGT GTA GAT TAT TCG ATA ACA TTG CGT CCG CAC GAA CCA ACA CCG GGC ATT TAC AAA
G R S S R S H H E R D R N D G R R R T
GGA CGC TCG TCA CGG TCG CAC CAC GAA AGA GAT CGA AAT GAC GGT CGT CGA CGA ACA
P P S P Y R E R R R Y S R S R S Y S
CCA CCA TCA CCG TAC CGA GAA CGT CGC CGT CGT TAC AGT AGA TCA CGT TCG TAT TCA
R E R K -
CGT GAA CGG AAG TAA TTT TTC GCG GGC ACT TTT CTT TTC TCC TTT CTT TTT TCT TCT
ACG ATA ATT TTG TAA GAT TTA GTT TGT TAA AAA GCA CAA AAG ATA GCG GGA CCA CAA
AAT TGT GGT TGT TTG TTC AAT TGG TTT ACC GAA AAT TAA AAA GAA AAA AGA GAA TTC
GGT ATG CCG AAG GTG AGA CTA ATC CGG GAC ATT TGT GGA ATT GTG TTG AAT CTT TAA
TTG TTT TTA AGA ATT TTC GCC ATG TCT TTT GAC CAT TTT CCT TTA TTA CTA TTG ATA
GAA GAC ATT TTA CTA TGA AGA ATT ATA GAC TTG CCA GTT AAC TGC TTA CCA TGA AGC
AGA TGT ATG AAA TTT TTC TTG GTT TTA TTT TAC AAA AGC AAA TTT CCG TAA ATC GGA
AAT CAA AGT CGC CTT AAA CAA TTC GAA ACA CCA AAA GCA TGG CAA ATT AGA TGA CAC
ACA GAC ATT TAA ATT AAT TGA AAT TAA AAC AAA CTG AGG ACA CCA TTC AAA ATG GTG
GAT ACT TTT AGG CTT AAG ACG CGG TGT TTT TCC TTC TTT TCT TTT TTT TGG TTT GGA
GCG TTC AAC TGA AAC TAT AAA TTG TGA TTA AAA ATT GAT TTG CAT TAA TGA AAT TTT
GTT GGA TAA GGC GAG GAA CAT TCA ATT AAT TAA TTA AAA GAA AAT AAA TTA ATT TCG
TTT GGA TAT ACA CTC TCA CAC TAC CTA ACA AAT TGT TTA TTT TTT TTT GGA ACT GAT
TTT CGT TTG TAA TTA GCA CAT TAA AAA GAA AAC TTA GTT GTA ATT TGA TTT AAA AAC
AAA ATA TTG AAA TGA AAC AAC CCA CGA TCG ATT CGA TCA AAA AAA TAA TTT AGA ATT
AAG ACA ACA CTA AAA TAT TAT AGA GAG AGA TGA TTG CAA TAA AAT GTG CGA GTA TTT
TTG AAA ATC 3'

```

**Secuencia de nucleótidos correspondiente al transcrito *Sotra2-251*.** En verde se indica la secuencia de aminoácidos de la proteína putativa que codifica, SoTra2-251.

```

5' G TGA AGC TGT GTC CTA TTG TAC AAT TGC AAT TAT AAA
AGC GAG CCA AGA ACC GCT ATA AAT ACA TCT CCA GTG TCA TTA ATC GAA TAC AGA ATA

M S R R M S R S R S R S Y E V H R R P
ATG TCT CGA AGG ATG TCA CGT TCG AGA AGT CGG AGT TAC GAA GTT CAC AGA CGA CCT
I E P P S P P S V G R K Y R D E S Y S
ATA GAA CCT CCT TCG CCT CCA TCC GTC GGC AGG AAG TAC AGA GAT GAG TCG TAT TCG
P Y K R R S D R R R S Y S R S P Y N D
CCG TAT AAG CGT CGC AGC GAT CGT CGC AGA TCC TAC TCC AGA AGC CCG TAC AAT GAT
R H R H V G T R E K P R K S R C L G V
CGC CAC AGA CAT GTG GGC ACA CGT GAA AAG CCG AGA AAG AGC CGT TGC CTT GGT GTG
F G L S V Y T T K D Q I V Q I F S K Y
TTT GGT CTA AGT GTC TAC ACC ACG AAA GAT CAA ATC GTG CAG ATA TTC TCC AAG TAT
G P I E H V Q V V V D A A T G R S R G
GGT CCA ATT GAA CAT GTT CAA GTG GTT GTC GAT GCG GCG ACA GGG CGT TCG CGT GGC
F C F V Y F D R T D D A T V A R D H C
TTT TGT TTC GTC TAC TTT GAT CGC ACC GAT GAC GCT ACC GTT GCC AGA GAC CAT TGC
T G M E I D N R R I R V D Y S I T L R
ACC GGC ATG GAA ATC GAT AAC CGT CGC ATT CGT GTA GAT TAT TCG ATA ACA TTG CGT
P H E P T P G I Y K G R S S R S H H E
CCG CAC GAA CCA ACA CCG GGC ATT TAC AAA GGA CGC TCG TCA CGG TCG CAC CAC GAA
R D R N D G R R R T P P S P Y R E R R
AGA GAT CGA AAT GAC GGT CGT CGA CGA ACA CCA CCA TCA CCG TAC CGA GAA CGT CGC
R R Y S R S R S Y S R E R K -
CGT CGT TAC AGT AGA TCA CGT TCG TAT TCA CGT GAA CGG AAG TAA TTT TTC GCG GGC
ACT TTT CTT TTC TCC TTT CTT TTT TCT TCT ACG ATA ATT TTG TAA GAT TTA GTT TGT
TAA AAA GCA CAA AAG ATA GCG GGA CCA CAA AAT TGT GGT TGT TTG TTC AAT TGG TTT
ACC GAA AAT TAA AAA GAA AAA AGA GAA TTC GGT ATG CCG AAG GTG AGA CTA ATC CGG
GAC ATT TGT GGA ATT GTG TTG AAT CTT TAA TTG TTT TTA AGA ATT TTC GCC ATG TCT
TTT GAC CAT TTT CCT TTA TTA CTA TTG ATA GAA GAC ATT TTA CTA TGA AGA ATT ATA
GAC TTG CCA GTT AAC TGC TTA CCA TGA AGC 3'

```

**Secuencia de nucleótidos correspondiente al transcrito *Sotra2-204*.** En verde se indica la secuencia de aminoácidos de la proteína putativa que codifica, SoTra2-204.

```

5'G TGA AGC TGT GTC CTA TTG TAC AAT TGC AAT TAT AAA
AGC GAG CCA AGA ACC GCT ATA AAT ACA TCT CCA GTG TCA TTA ATC GAA TAC AGA ATA
M S R R M S R S R S R S Y E V H R R P
ATG TCT CGA AGG ATG TCA CGT TCG AGA AGT CGG AGT TAC GAA GTT CAC AGA CGA CCT
I E P P S P P S V G R K Y R D D D S R
ATA GAA CCT CCT TCG CCT CCA TCC GTC GGC AGG AAG TAC AGA GAT GAC GAC AGC CGC
S G S R E V T L R K K R R M S Y D R S
AGT GGT TCT CGA GAA GTG ACC CTT CGA AAG AAG CGA AGA ATG TCA TAT GAT CGG TCA
R S R S R S R S M G R Y A H R S S R Y
AGA TCA CGA TCT CGT TCA CGA TCA ATG GGA CGA TAC GCT CAC AGA TCA AGC AGA TAT
V R S R S R K K Y P K E K K Y F Q S T
GTG CGA AGT CGA TCT CGG AAG AAA TAT CCA AAA GAG AAA AAA TAT TTT CAA AGC ACG
G N T S I N L I K K S Q P E I Y N S T
GGA AAC ACT TCA ATC AAT TTA ATA AAA AAA AGT CAA CCC GAA ATC TAT AAT TCA ACT
L T K T N S T T K T D S I T R T -
TTA ACC AAA ACC AAT TCT ACA ACA AAA ACC GAC TCA ATT ACA CGA ACA TAG TCC CCT
CAA GAA ATA TCA TTT TTT GTG ATT TGA AAT TCA GCA ACC TAA CCC AAG GCC TGA TGA
TAA AAC AAT ACA AAC AAA CGA TCG AAG AAT GTG TTT GCG CTC TTG TTC TTT TGA TGA
AGG ATC AAA CCC AAC GAA TTA TAC ACT CGA AGC AAC CAC AAG CAT AAC AAA AAA TTA
CCG AAA CGA GCA CAC GCA TTA CAA ACT GAC ATC GCC GCC GTC CTT CGA TCA TAA TTT
GTG AAA ACG TGT CAA ACC ATC GAA TAA ATC CAA ACA AAA ATC TTT TCT TCC ACA AAA
AAA GGT CGT ATT CGC CGT ATA AGC GTC GCA GCG ATC GTC GCA GAT CCT ACT CCA GAA
GCC CGT ACA ATG ATC GCC ACA GAC ATG TGG GCA CAC GTG AAA AGC CGA GAA AGA GCC
GTT GCC TTG GTG TGT TTG GTC TAA GTG TCT ACA CCA CGA AAG ATC AAA TCG TGC AGA
TAT TCT CCA AGT ATG GTC CAA TTG AAC ATG TTC AAG TGG TTG TCG ATG CGG CGA CAG
GGC GTT CGC GTG GCT TTT GTT TCG TCT ACT TTG ATC GCA CCG ATG ACG CTA CCG TTG
CCA GAG ACC ATT GCA CCG GCA TGG AAA TCG ATA ACC GTC GCA TTC GTG TAG ATT ATT
CGA TAA CAT TGC GTC CGC ACG AAC CAA CAC CGG GCA TTT ACA AAG GAC GCT CGT CAC
GGT CGC ACC ACG AAA GAG ATC GAA ATG ACG GTC GTC GAC GAA CAC CAC CAT CAC CGT
ACC GAG AAC GTC GCC GTC GTT ACA GTA GAT CAC GTT CGT ATT CAC GTG AAC GGA AGT
AAT TTT TCG CGG GCA CTT TTC TTT TCT CCT TTC TTT TTT CTT CTA CGA TAA TTT TGT
AAG ATT TAG TTT GTT AAA AAG CAC AAA AGA TAG CGG GAC CAC AAA ATT GTG GTT GTT
TGT TCA ATT GGT TTA CCG AAA ATT AAA AAG AAA AAA GAG AAT TCG GTA TGC CGA AGG
TGA GAC TAA TCC GGG ACA TTT GTG GAA TTG TGT TGA ATC TTT AAT TGT TTT TAA GAA
TTT TCG CCA TGT CTT TTG ACC ATT TTC CTT TAT TAC TAT TGA TAG AAG ACA TTT TAC
TAT GAA GAA TTA TAG ACT TGC CAG TTA ACT GCT TAC CAT GAA GC 3'

```

**Secuencia de nucleótidos correspondiente al transcrito *Sotra2-130*.** En verde se indica la secuencia de aminoácidos de la proteína putativa que codifica, SoTra2-130.



5'GTGAAGCT	CTGCTATTGC	ACAATTGTGA	TTGCAACGGC	GAGCTAGAAA	CCGCTATAAA	TACATCTCCA
GTGTCATTAA	TCGAATGCGA	ATCAAATGTC	TCGAATGGTA	AGCATAAACC	AGATCGTTCC	TTCAATTCTT
TTTTTTGTCA	AGCCTGTCAA	CCTATAAATA	TAATTCTCAA	ATTCAAATTA	TAATTGAAA	ATAAAGCAAT
CATAGACATC	GGGTTGTTTT	ATGAAAAACC	ATAGCAACAC	GGCTTTTGTT	CATCCTTTGC	TTAAATTTGT
TTTTTTCATT	GATCAAACGA	AAGCCACGCT	TGAAAGTGAT	AAGTGCCGA	AACTCAAGCT	GGTGTTTATT
TTCAGAAAAC	CAATCTTCTA	GATTTAAATA	ATGAGTTTGC	TTGTAATTGC	CATTTTTTGT	TCGCGTGGTG
TTAGCGGTAC	AAATTCGATA	AATTTGTTTT	TTTTGTTTTA	GATGTCTCGT	TCGAGAAGTC	GGAGTTATGA
AGTTCACAAA	AGACCTATTG	AACCTCCGTC	GCCACCATCC	GTCGGCAGGA	AGTACAGAGA	TGAGTAAAGTA
ACATTAATTC	AGTTGGGGTG	GGTACCTTTT	TATGGTGACT	TGATTTGCAC	TCTGAGTCAC	ATATTTTTTAC
ATAAATACTG	TTCACAAAGC	TAATAATGCA	TCCGGACGTT	TAAACAGACA	GTACATTTTG	TTCTAGCCAA
TAATGATGTG	AAACGTTACG	GGTGAGGTTT	AATGTTATTG	TAGTACATTC	GACAGTCTGA	TTATGCAATT
TTCAACGTTT	CTAACGTACG	CTTTGCGTTT	CATTCGTTGA	TAGCGACTAG	GGTAACGGAA	CCACGAGCCC
TGAAAGAGAA	AAGCGAACGG	TACAGCTCTG	TGCATGAGAA	AGAGACAAAG	CAGGGGCTCG	TAATAGAAAT
AAAGGGATCT	CGTAACCGGT	ACCGTTACTC	TAGTACTTTG	TTAATCATCA	CAATTCCAAT	GCCAATCTTC
CGTTTCGGCG	CCGACGGCCA	CTGGTTTTCT	AAATTTAAAT	GTTTAATCTG	ATGCAACGAC	ACAAATCCAC
TACTTATATG	CCGTTGCAAG	TCGGATAGAC	GGTCTTATTG	ATTGGATGAT	GCTCAACGGC	TCTTCTGGCT
GAACAAAAAT	TGCATTTCAA	TTTCACTTGG	ATGCATCATA	AATATTAACC	TTAGCAACCA	ACTTCGAAAC
ATATAATTGA	TTAATAAGAG	TTGACTGAGC	ACTGCCTTAG	AACTCGGAAC	AGATATCAGG	TTTACCTGTG
CCTCATATGA	AACCATCAAT	TTTGAATTAC	CTGAAGGCTG	ACATGATTTG	ACTGTAATCG	GAGCTTCAAC
ACAGATGAAG	TTCTTTTTCT	ACGAACAGGG	ATTCGACAAG	CTCGATAAAG	CTCTTCGGGC	GTTGTATGAA
GAAAATTAGT	AAAAAGTCCT	TCATTGCCCT	TCAATCATAT	GAAAGAAGAA	GAATGCCCTT	CTCGCCCGAA
TGCCGCTAGT	CTTCTCGGG	CCTGAGCAAC	TAGTTCGATA	AATCCTTCAG	GTGCTGCATG	AAGAAAATTT
GTAAATCAA	AATGCCCTTA	GGTCTAGTTC	GGGCCTGACA	ATACGTGTAC	TCGTCGTTTT	TCTATCACCT
GACCAAGTCA	GGCTGACCTG	TACTTTCAAG	TCAAAC TGAC	GTCTAGGCGT	CGAGGTTTAT	GACACAGTTG
AATTTTAACT	TTGTTAGTCT	CGAGAGCATT	CGAATCATCC	CCAGACACTT	CGACATAAAA	TATATGTGAA
TAGCGTGAAT	AGTGTCACAC	CCGTGATTCT	AATCTTTCTC	CATTTCAGAGT	CTTTGAAACT	AGAGCTCCAG
TGACAATATT	TCTTTAAAC	TCCACAATCG	ACAACCATAG	TAAGGCAAAA	TGTATTTTTT	CAGCGACAGC
CGCAGCGGTT	CTCGAGAAGT	GACACTTCGA	AAGAAGCGAA	AAATGTCGTA	TGATCGATTG	CGTTCACGAT
CTATGGGTCG	ACACGGTCAC	AGATCAAGCA	GATATGAGCG	GAGCCGATCT	CGGTAAGTTA	CGACTCCGAA
AGCATTGTTT	AGTGAAAGCA	CACATAAAAT	TTGAAAATTA	TTTTTTCATT	CCCATCAGCT	GTGAAGCGTA
TATATAGATG	TAACGATTTT	TTTCTTATC	TCTTTTTTTC	TGCTCTGAAG	GAAGAAATGT	CCTAAGAAAA
AAAGAACTTT	TCAAACACAC	GGGAAACACT	TCAATCAATA	TATAAAAAAG	TGACTTTGAA	AAATAGTTCA
ACTTTACCAA	AATCATTCGA	CAACAAAAAT	CGACTCAAAT	TACACACATT	TCCCTCAAGA	ATCTATCTTT
TGTGAATTGA	CATTTCAGCA	TCTAACCGCG	ACCTGATTAG	TTGATCGAAG	AATATGTTTG	CGCTCTATTG
TTCTTGACGG	GGGATCCAAA	TCCAACGAAT	TATACTCCAA	TCAAGCACAA	AACAAAACCA	AACATACCGA
AACGAAATGA	ACGCATTGCG	AACCATCGAC	GCCATTCTTC	GATCATAATT	CCGTAAAACG	TAGCCAACGA
ATAAACCAAA	AACAAAAATC	TTTTCTTTTC	CAACAAAAGG	TCGTTCTCGC	CGCATAAGCG	TCGCAGTGAT
CGTCGCCGAT	CTTACTCCAG	AAGCCCGTAC	AACGATCGCT	GGAGACACGT	GGGTACCCGG	GTAAGTGTA
GGATTTGGTA	CACACGGCGA	GAGTGCAACT	ATTTTGATTT	CCACTTTTTT	TTTAATAGGA	AAAACCGAAG
AAAAGCCGTT	GCCTTGAGGT	GTTTGGTCTG	AGTGTCTACA	CCACGAAAGA	TCAAATCATA	CAGATATTTT
CAAAGTACGG	ATCGATCGAG	CACGTTTCAGG	TTGTGGTCTG	TGCGTCCGGT	AGTTAATCAA	AATTTTTTCAT
TTAATTACCG	ATAGCCACAG	CACAGTACGA	ATAGTTGAAA	GCATAAAATT	ATCTCGTAAA	GTGCAACGTG
GCTAGATTGG	AGACCATTTG	AGAGAATTAG	AACGAGATTG	CCTGTAGCCT	TGGAATCTCA	TTAAAGGCCT
CACAAAACAA	ACTAGATTTT	GAACGATCAG	GCGGTTAGTT	TGCCACCGGC	AGACATGCAA	GAAGTGTTTT
CTTTTAATCA	AGTTAAAACT	AGCGGTTTCC	GATGCTTTTC	GTTCTATCAT	CCTCACTGTA	CTACTGGAAG
TTTTGTGTTT	ACTTCTCGAT	GAGACATTTA	AGTAATTTTG	AGCCTTATCA	TAATTGCCTT	AAGAACCCTA

```

TTTCTCGATC  GATTTTGTGA  GTCAAACGAA  CTCAGAATTC  GCATTACGAC  GATTGTCCAG  AGTCCCGCTT
TCAGGCCCTA  CATTCTTAGG  AAGGACCACG  TAGCCGAGCG  CCTGTTAGGT  CTGATGTATA  TGTAACATGT
TTCCAGTGTC  GGTCCTAGTT  AAGCATAAGA  CGTGCAATGC  GCGTATTCAG  TTGAAAGTCG  TCGAAGTACG
GTTGCCATAC  GAAGAGCGTT  GTAATGTTAG  CACTACCATG  TTACAGTATG  TTACGGTGAC  AACTAACTCA
ACTGTCGTTA  GGTCAAAACA  ACTGTGGGAA  ATTGATCGTT  CGGATCGTTC  ATTAATTTAA  TGTCTCCAC
GCAAACACT  TCCACCTTCC  ATGAACTGCA  ATGCCTCAAC  CGTTCTAATT  GTTGTTTTTT  AACAAATTC
CAGACTGGCC  GTTCCCGTGG  CTTCTGTTTC  GTCTACTTTG  ATCGCACCAG  AGATGCAACC  GTTGCCAGAG
ACCACTGCAC  TGGCATGGAA  ATCGATAATC  GTCGCATTCT  GGTGACTAT  TCAATAACGT  TCGGGCCGCA
CGGACCAACA  CCGGGCATTT  ACAAAGGTCG  TTCGACTCGG  TCGCACTATG  AAAGAGATCG  CAGTGACGGT
CGTCGCCGGT  CACCGTCACC  CTATCGAGAG  CGTCGTCGAC  GCTACAGCAG  ATCGCGCTCG  TATTCACGTG
AACGGAAATA  ATTTTCCCAT  TTTTCCTTGA  AAACCTTTTC  TTCTTTTCTT  TTTCTTTGCT  AATTTTAAAG
GTTTTGATTT  GTTCAAAGCA  CACAAAAAAA  AAACGAAACG  GACCACATTG  ATCTCTTTTG  TTCGTTTGCC
TTACCGATGT  GTTGGTATTA  TCTGAATGGT  AGTATATTAA  TCCCGGGACA  TTTGTGGAAT  TGTGTTGAAT
CTTTATTGTT  TTAAGATCTT  TTTTTCGTCG  TTATGAATTG  CCATGCTTTT  TTGACCAGCC  TTTTCTATG
ATAAGGACAT  TAAGTACTTT  AAAAATTACA  GACTTTCCGA  TGTTACCACA  AAAAAT 3'

```

**Secuencia de nucleótidos correspondiente al gen *tra2* de *S.coprophila*.** Los exones aparecen en color verde y los intrones en color rojo. La secuencia subrayada corresponde al intrón que se retiene parcialmente en alguna de las formas (E4a). El recuadro negro señala la secuencia ISS. En azul se indican las regiones 5' y 3' UTR.

```

5'  GTG AAG CTC TGC TAT TGC ACA ATT GTG ATT GCA ACG GCG AGC TAG AAA CCG CTA
TAA ATA CAT CTC CAG TGT CAT TAA TCG AAT GCG AAT CAA ATG TCT CGA ATG GTA AGC
ATA AAC CAG ATC GTT CCT TCA ATT CTT TTT TTT GTC AAG CCT GTC AAC CTA TAA ATA
TAA TTC TCA AAT TCA AAT TAT AAT TCG AAA ATA AAG CAA TCA TAG ACA TCG GGT TGT
TTT ATG AAA AAC CAT AGC AAC ACG GCT TTT GTT CAT CCT TTG CTT AAA TTT GTT TTT
TTC ATT GAT CAA ACG AAA GCC ACG CTT GAA AGT GAT AAG TGG CCG AAA CTC AAG CTG
GTG TTT ATT TTC AGA AAA CCA ATC TTC TAG ATT TAA ATA

M  S  L  L  M  S  R  S  R  S  R  S  Y  E  V  H  K  R  P
ATG AGT TTG CTT ATG TCT CGT TCG AGA AGT CGG AGT TAT GAA GTT CAC AAA AGA CCT

I  E  P  P  S  P  P  S  V  G  R  K  Y  R  D  D  D  S  R
ATT GAA CCT CCG TCG CCA CCA TCC GTC GGC AGG AAG TAC AGA GAT GAC GAC AGC CGC

S  G  S  R  E  V  T  L  R  K  K  R  K  M  S  Y  D  R  L
AGC GGT TCT CGA GAA GTG ACA CTT CGA AAG AAG CGA AAA ATG TCG TAT GAT CGA TTG

R  S  R  S  M  G  R  H  G  H  R  S  S  R  Y  E  R  S  R
CGT TCA CGA TCT ATG GGT CGA CAC GGT CAC AGA TCA AGC AGA TAT GAG CGG AGC CGA

S  R  S  F  S  P  H  K  R  R  S  D  R  R  R  S  Y  S  R
TCT CGG TCG TTC TCG CCG CAT AAG CGT CGC AGT GAT CGT CGC CGA TCT TAC TCC AGA

S  P  Y  N  D  R  W  R  H  V  G  T  R  E  K  P  K  K  S
AGC CCG TAC AAC GAT CGC TGG AGA CAC GTG GGT ACC CGG GAA AAA CCG AAG AAA AGC

R  C  L  G  V  F  G  L  S  V  Y  T  T  K  D  Q  I  I  Q
CGT TGC CTT GGA GTG TTT GGT CTG AGT GTC TAC ACC ACG AAA GAT CAA ATC ATA CAG

I  F  S  K  Y  G  S  I  E  H  V  Q  V  V  V  D  A  S  T
ATA TTT TCA AAG TAC GGA TCG ATC GAG CAC GTT CAG GTT GTG GTC GAT GCG TCG ACT

G  R  S  R  G  F  C  F  V  Y  F  D  R  T  E  D  A  T  V
GGC CGT TCC CGT GGC TTC TGT TTC GTC TAC TTT GAT CGC ACC GAA GAT GCA ACC GTT

A  R  D  H  C  T  G  M  E  I  D  N  R  R  I  R  V  D  Y
GCC AGA GAC CAC TGC ACT GGC ATG GAA ATC GAT AAT CGT CGC ATT CGG GTT GAC TAT

S  I  T  L  R  P  H  G  P  T  P  G  I  Y  K  G  R  S  T
TCA ATA ACG TTG CGG CCG CAC GGA CCA ACA CCG GGC ATT TAC AAA GGT CGT TCG ACT

R  S  H  Y  E  R  D  R  S  D  G  R  R  R  S  P  S  P  Y
CGG TCG CAC TAT GAA AGA GAT CGC AGT GAC GGT CGT CGC CGG TCA CCG TCA CCC TAT

R  E  R  R  R  R  Y  S  R  S  R  S  Y  S  R  E  R  K  -
CGA GAG CGT CGT CGA CGC TAC AGC AGA TCG CGC TCG TAT TCA CGT GAA CGG AAA TAA

TTT TCC CAT TTT TCC TTG AAA ACT TTT TCT TCT TTT CTT TTT CTT TGC TAA TTT TTA
AGG TTT TGA TTT GTT CAA AGC ACA CAA AAA AAA AAC GAA ACG GAC CAC ATT GAT CTC
TTT TGT TCG TTT GCC TTA CCG ATG TGT TGG TAT TAT CTG AAT GGT AGT ATA TTA ATC
CCG GGA CAT TTG TGG AAT TGT GTT GAA TCT TTA TTG TTT TAA GAT CTT TTT TTT TCG
TTT ATG AAT TGC CAT GCT TTT TTG ACC AGC CTT TTT CTA TGA TAA GGA CAT TAA GTA
CTT TAA AAA TTA CAG ACT TTC CGA TGT TAC CAC AAA AAA T 3'

```

**Secuencia de nucleótidos correspondiente al transcrito *ScTra2-246*.** En verde se indica la secuencia de aminoácidos de la proteína putativa que codifica, ScTra2-246.

```

5' GTG AAG CTC TGC TAT TGC
ACA ATT GTG ATT GCA ACG GCG AGC TAG AAA CCG CTA TAA ATA CAT CTC CAG TGT CAT TAA
TCG AAT GCG AAT CAA ATG TCT CGA ATG GTA AGC ATA AAC CAG ATC GTT CCT TCA ATT CTT
TTT TTT GTC AAG CCT GTC AAC CTA TAA ATA TAA TTC TCA AAT TCA AAT TAT AAT TCG AAA
ATA AAG CAA TCA TAG ACA TCG GGT TGT TTT ATG AAA AAC CAT AGC AAC ACG GCT TTT GTT
CAT CCT TTG CTT AAA TTT GTT TTT TTC ATT GAT CAA ACG AAA GCC ACG CTT GAA AGT GAT
AAG TGG CCG AAA CTC AAG CTG GTG TTT ATT TTC AGA AAA CCA ATC TTC TAG ATT TAA ATA
M S L L M S R S R S R S Y E V H K R P I
ATG AGT TTG CTT ATG TCT CGT TCG AGA AGT CGG AGT TAT GAA GTT CAC AAA AGA CCT ATT
E P P S P P S V G R K Y R D E S F S P H
GAA CCT CCG TCG CCA CCA TCC GTC GGC AGG AAG TAC AGA GAT GAG TCG TTC TCG CCG CAT
K R R S D R R R S Y S R S P Y N D R W R
AAG CGT CGC AGT GAT CGT CGC CGA TCT TAC TCC AGA AGC CCG TAC AAC GAT CGC TGG AGA
H V G T R E K P K K S R C L G V F G L S
CAC GTG GGT ACC CGG GAA AAA CCG AAG AAA AGC CGT TGC CTT GGA GTG TTT GGT CTG AGT
V Y T T K D Q I I Q I F S K Y G S I E H
GTC TAC ACC ACG AAA GAT CAA ATC ATA CAG ATA TTT TCA AAG TAC GGA TCG ATC GAG CAC
V Q V V V D A S T G R S R G F C F V Y F
GTT CAG GTT GTG GTC GAT GCG TCG ACT GGC CGT TCC CGT GGC TTC TGT TTC GTC TAC TTT
D R T E D A T V A R D H C T G M E I D N
GAT CGC ACC GAA GAT GCA ACC GTT GCC AGA GAC CAC TGC ACT GGC ATG GAA ATC GAT AAT
R R I R V D Y S I T L R P H G P T P G I
CGT CGC ATT CGG GTT GAC TAT TCA ATA ACG TTG CGG CCG CAC GGA CCA ACA CCG GGC ATT
Y K G R S T R S H Y E R D R S D G R R R
TAC AAA GGT CGT TCG ACT CGG TCG CAC TAT GAA AGA GAT CGC AGT GAC GGT CGT CGC CGG
S P S P Y R E R R R R Y S R S R S Y S R
TCA CCG TCA CCC TAT CGA GAG CGT CGT CGA CGC TAC AGC AGA TCG CGC TCG TAT TCA CGT
E R K -
GAA CGG AAA TAA TTT TCC CAT TTT TCC TTG AAA ACT TTT TCT TCT TTT CTT TTT CTT TGC
TAA TTT TTA AGG TTT TGA TTT GTT CAA AGC ACA CAA AAA AAA AAC GAA ACG GAC CAC ATT
GAT CTC TTT TGT TCG TTT GCC TTA CCG ATG TGT TGG TAT TAT CTG AAT GGT AGT ATA TTA
ATC CCG GGA CAT TTG TGG AAT TGT GTT GAA TCT TTA TTG TTT TAA GAT CTT TTT TTT TCG
TTT ATG AAT TGC CAT GCT TTT TTG ACC AGC CTT TTT CTA TGA TAA GGA CAT TAA GTA CTT
TAA AAA TTA CAG ACT TTC CGA TGT TAC CAC AAA AAA T 3'

```

**Secuencia de nucleótidos correspondiente al transcrito *ScTra2-203*.** En verde se indica la secuencia de aminoácidos de la proteína putativa que codifica, ScTra2-203.

```

5' GTG AAG CTC TGC TAT TGC
ACA ATT GTG ATT GCA ACG GCG AGC TAG AAA CCG CTA TAA ATA CAT CTC CAG TGT CAT TAA
TCG AAT GCG AAT CAA ATG TCT CGA ATG GTA AGC ATA AAC CAG ATC GTT CCT TCA ATT CTT
TTT TTT GTC AAG CCT GTC AAC CTA TAA ATA TAA TTC TCA AAT TCA AAT TAT AAT TCG AAA
ATA AAG CAA TCA TAG ACA TCG GGT TGT TTT ATG AAA AAC CAT AGC AAC ACG GCT TTT GTT
CAT CCT TTG CTT AAA TTT GTT TTT TTC ATT GAT CAA ACG AAA GCC ACG CTT GAA AGT GAT
AAG TGG CCG AAA CTC AAG CTG GTG TTT ATT TTC AGA AAA CCA ATC TTC TAG ATT TAA ATA
M S L L M S R S R S R S Y E V H K R P I
ATG AGT TTG CTT ATG TCT CGT TCG AGA AGT CGG AGT TAT GAA GTT CAC AAA AGA CCT ATT
E P P S P P S V G R K Y R D D D S R S G
GAA CCT CCG TCG CCA CCA TCC GTC GGC AGG AAG TAC AGA GAT GAC GAC AGC CGC AGC GGT
S R E V T L R K K R K M S Y D R L R S R
TCT CGA GAA GTG ACA CTT CGA AAG AAG CGA AAA ATG TCG TAT GAT CGA TTG CGT TCA CGA
S M G R H G H R S S R Y E R S R S R K K
TCT ATG GGT CGA CAC GGT CAC AGA TCA AGC AGA TAT GAG CGG AGC CGA TCT CGG AAG AAA
C P K K K R T F Q T H G K H F N Q Y I K
TGT CCA AAG AAA AAA AGA ACT TTT CAA ACA CAC GGG AAA CAC TTC AAT CAA TAT ATA AAA
K S T -
AAG TCG ACT TGA AAA ATA GTT CAA CTT TAC CAA AAT CAT TCG ACA ACA AAA ATC GAC TCA
AAT TAC ACA CAT TTC CCT CAA GAA TCT ATC TTT TGT GAA TTG ACA TTC AGC AAT CTA ACC
GCG ACC TGA TTA GTT GAT CGA AGA ATA TGT TTG CGC TCT ATT GTT CTT GAC GGG GGA TCC
AAA TCC AAC GAA TTA TAC TCC AAT CAA GCA CAA AAC AAA ACC AAA CAT ACC GAA ACG AAA
TGA ACG CAT TGC GAA CCA TCG ACG CCA TTC TTC GAT CAT AAT TCC GTA AAA CGT AGC CAA
CGA ATA AAC CAA AAA CAA AAA TCT TTT CTT TCA CAA CAA AAG GTC GTT CTC GCC GCA TAA
GCG TCG CAG TGA TCG TCG CCG ATC TTA CTC CAG AAG CCC GTA CAA CGA TCG CTG GAG ACA
CGT GGG TAC CCG GGA AAA ACC GAA GAA AAG CCG TTG CCT TGG AGT GTT TGG TCT GAG TGT
CTA CAC CAC GAA AGA TCA AAT CAT ACA GAT ATT TTC AAA GTA CGG ATC GAT CGA GCA CGT
TCA GGT TGT GGT CGA TGC GTC GAC TGG CCG TTC CCG TGG CTT CTG TTT CGT CTA CTT TGA
TCG CAC CGA AGA TGC AAC CGT TGC CAG AGA CCA CTG CAC TGG CAT GGA AAT CGA TAA TCG
TCG CAT TCG GGT TGA CTA TTC AAT AAC GTT GCG GCC GCA CGG ACC AAC ACC GGG CAT TTA
CAA AGG TCG TTC GAC TCG GTC GCA CTA TGA AAG AGA TCG CAG TGA CGG TCG TCG CCG GTC
ACC GTC ACC CTA TCG AGA GCG TCG TCG ACG CTA CAG CAG ATC GCG CTC GTA TTC ACG TGA
ACG GAA ATA ATT TTC CCA TTT TTC CTT GAA AAC TTT TTC TTC TTT TCT TTT TCT TTG CTA
ATT TTT AAG GTT TTG ATT TGT TCA AAG CAC ACA AAA AAA AAA CGA AAC GGA CCA CAT TGA
TCT CTT TTG TTC GTT TGC CTT ACC GAT GTG TTG GTA TTA TCT GAA TGG TAG TAT ATT AAT
CCC GGG ACA TTT GTG GAA TTG TGT TGA ATC TTT ATT GTT TTA AGA TCT TTT TTT TTC GTT

```

TAT GAA TTG CCA TGC TTT TTT GAC CAG CCT TTT TCT ATG ATA AGG ACA TTA AGT ACT TTA  
AAA ATT ACA GAC TTT CCG ATG TTA CCA CAA AAA AT 3´

**Secuencia de nucleótidos correspondiente al transcrito *ScTra2-99*.** En verde se indica la secuencia de aminoácidos de la proteína putativa que codifica, ScTra2-99.

```

5' GTG AAG CTC TGC TAT TGC
ACA ATT GTG ATT GCA ACG GCG AGC TAG AAA CCG CTA TAA ATA CAT CTC CAG TGT CAT TAA
TCG AAT GCG AAT CAA ATG TCT CGA ATG GTA AGC ATA AAC CAG ATC GTT CCT TCA ATT CTT
TTT TTT GTC AAG CCT GTC AAC CTA TAA ATA TAA TTC TCA AAT TCA AAT TAT AAT TCG AAA
ATA AAG CAA TCA TAG ACA TCG GGT TGT TTT ATG AAA AAC CAT AGC AAC ACG GCT TTT GTT
CAT CCT TTG CTT AAA TTT GTT TTT TTC ATT GAT CAA ACG AAA GCC ACG CTT GAA AGT GAT
AAG TGG CCG AAA CTC AAG CTG GTG TTT ATT TTC AGA AAA CCA ATC TTC TAG ATT TAA ATA
M   S   L   L   M   S   R   S   R   S   R   S   Y   E   V   H   K   R   P   I
ATG AGT TTG CTT ATG TCT CGT TCG AGA AGT CGG AGT TAT GAA GTT CAC AAA AGA CCT ATT
E   P   P   S   P   P   S   V   G   R   K   Y   R   D   E   L   A   V   P   V
GAA CCT CCG TCG CCA CCA TCC GTC GGC AGG AAG TAC AGA GAT GAA CTG GCC GTT CCC GTG
A   S   V   S   S   T   L   I   A   P   K   M   Q   P   L   P   E   T   T   A
GCT TCT GTT TCG TCT ACT TTG ATC GCA CCG AAG ATG CAA CCG TTG CCA GAG ACC ACT GCA
L   A   W   K   S   I   I   V   A   F   G   L   T   I   Q   -
CTG GCA TGG AAA TCG ATA ATC GTC GCA TTC GGG TTG ACT ATT CAA TAA CGT TGC GGC CGC
ACG GAC CAA CAC CGG GCA TTT ACA AAG GTC GTT CGA CTC GGT CGC ACT ATG AAA GAG ATC
GCA GTG ACG GTC GTC GCC GGT CAC CGT CAC CCT ATC GAG AGC GTC GTC GAC GCT ACA GCA
GAT CGC GCT CGT ATT CAC GTG AAC GGA AAT AAT TTT CCC ATT TTT CCT TGA AAA CTT TTT
CTT CTT TTC TTT TTC TTT GCT AAT TTT TAA GGT TTT GAT TTG TTC AAA GCA CAC AAA AAA
AAA ACG AAA CGG ACC ACA TTG ATC TCT TTT GTT CGT TTG CCT TAC CGA TGT GTT GGT ATT
ATC TGA ATG GTA GTA TAT TAA TCC CGG GAC ATT TGT GGA ATT GTG TTG AAT CTT TAT TGT
TTT AAG ATC TTT TTT TTT CGT TTA TGA ATT GCC ATG CTT TTT TGA CCA GCC TTT TTC TAT
GAT AAG GAC ATT AAG TAC TTT AAA AAT TAC AGA CTT TCC GAT GTT ACC ACA AAA AAT 3'

```

**Secuencia de nucleótidos correspondiente al transcrito *ScTra2-75*.** En verde se indica la secuencia de aminoácidos de la proteína putativa que codifica, ScTra2-75.

